كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



إعداد

أ.د. محمد سويلم حامد – د. هاجر عبد المولي

الاحصاء الحيوي

Biostatistics



المحاضرة الأولى

موضوعات الفصل الأول

تعريف علم الإحصاء	أهمية علم الإحصاء	خطوات البحث أو التحليل الإحصائي	تقسيمات علم الإحصاء
إستخدامات علم الإحصاء	انواع البيانات الإحصائية	أساليب جمع البيانات الإحصائية	العينات
	عرض البيانات	التوزيعات التكراريه	

كليات الخليج
Gulf Colleges

المفاهيم العامة للإحصاء الحيوي

مدلول كلمة الإحصاء:

إن كلمة الإحصاء (Statistique أو Statistics) كانت تهدف في الماضي إلى العد ومصدرها "أحصى" عدَّ. حتى أطلق عليه بعلم العد (Le science du Comptage) وكلمة إحصاء مصدرها State وتعني الدولة أي كانت تستعمل كلمة (Statistics) في أوربا للدلالة على أعمال وحسابات الدولة في شؤون الحرب والضرائب والسكان والمواليد والوفيات والإنتاج الزراعي.... الخ وللإحصاء مدلولان:

1- المدلول الأول: قائم على استعماله كمرادف لكلمة بيانات – أرقام – أعداد. مثال: عدد الطلاب في كلية التجارة. عدد المرضى المصابين بمرض السيدا.

2- المدلول الثاني: قائم على أنه مجموعة من الأساليب الفنية أو الطرق العلمية التي تستخدم في معالجة البيانات الرقمية وتعميم النتائج المستخلصة على المجتمعات المدروسة.

كليات الخليج
Gulf Colleges

اهمية علم الاحصاء الحيوى :-

يحتل الاحصاء مكاناً بين العلوم لما له من استعمالات واسعة للوصول الى قرارات صائبة لوصف او تفسير الظواهر المختلفة في جميع العلوم وهو المستعمل من قبل للأفراد والجماعات المختلفة والدول على حد السواء ، واستخدم الاحصاء في مجالات كثيرة ونركز على اهمية علم الاحصاء في العلوم البيولوجية والطبية والصحة العامة والكيمياء .

في علم الاحياء (البيولوجي) :-تستخدم الطرق الاحصائية في دراسة الاجناس والفصائل المختلفة للحيوان والنبات ومعرفة خواص كل جنس بما يتميز عن غيره واختلاف مفردات الجنس الواحد في اية خاصية معينة من الناحية الاحصائية ، فمثلاً نرى الذكور في الجنس البشري اطول قامة من الاناث مع ان الذكور فيما بينهم يختلفون في الطول الى درجة ما وكذلك الاناث ، كل ذلك يتم عن طريق جمع البيانات وتبويبها ودراستها دراسة احصائية والخروج بنتائج من هذه الصفات .

في الطب والصيدلة يستخدم الاحصاء لدراسة العلاقة بين متغيرات كثيرة منها على سبيل المثال العلاقة بين العمر وضغط الدم وكذلك العلاقة بين الوراثة والبيئة وتأثيراتهما على تكوين الفرد .

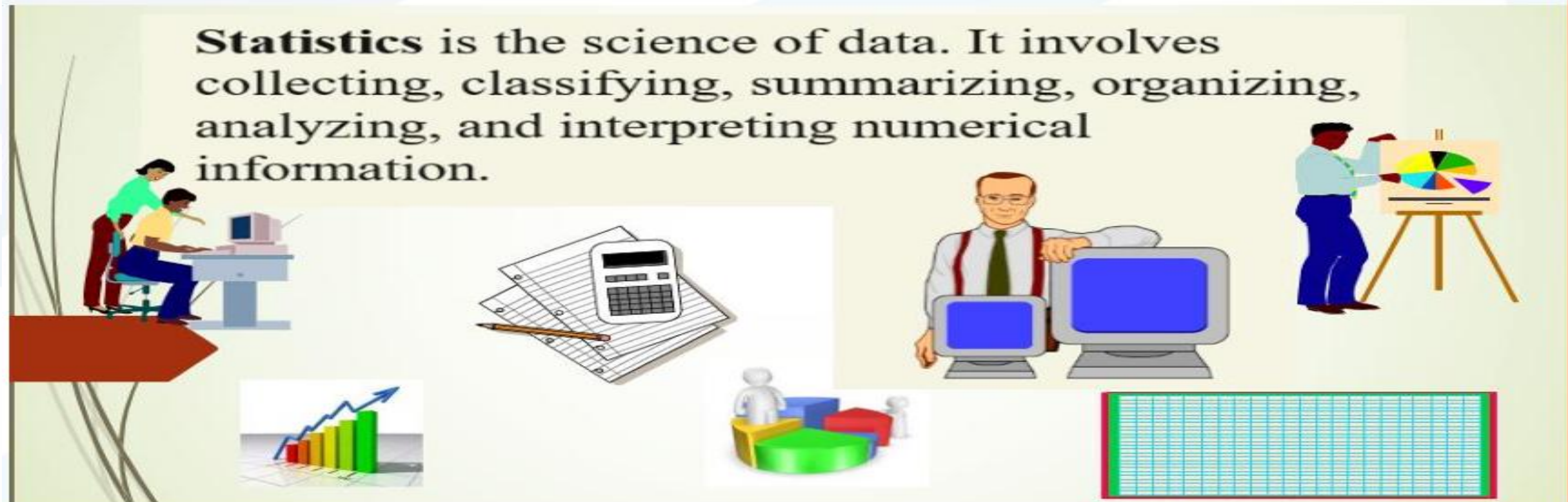
في الصحة العامة :- يستخدم الاحصاء لدراسة الامراض السارية والعقاقير الواجب استخدامها ونسبة زيادتها ونقصها في المجتمع وكذلك دراسة حالة المعوقين والوفيات ونسبة الزيادة في السكان .

* تعريف علم الإحصاء: Statistics

*علم الإحصاء: فرع من فروع المنهجية العلمية، فهو علم النظرية والأسلوب، ويختص بالطرق العلمية من: تجميع- تنظيم- تلخيص- عرض- تحليل- تفسير النتائج التي تم الحصول عليها، من المسوحات أو التجارب الإحصائية. بهدف الوصول إلى استدلالات محدّدة عن معالم المجتمع الإحصائي.

تعريف الاحصاء الحيوى (Biostatistics) .

يعني الاحصاء الحيوي اشياء مختلفة للأشخاص المختلفين فهو للعامة جداول واعداد عن البيانات الحياتية اما المعنى الاصطلاحي للإحصاء فهو رياضيات جمع البيانات للظواهر الطبية والصحية وتنظيمها وتحليلها وتفسيرها والتعميم من الخاص الى العام عن طريق استدلال خواص المجتمع من خواص العينة .



What Is Statistics?

1. **Collecting Data**
   e.g., Survey
2. **Presenting Data**
   e.g., Charts & Tables
3. **Characterizing Data**
   e.g., Average



Data Analysis



Why?



Decision-Making

. وباختصار الإحصاء علم وفن.

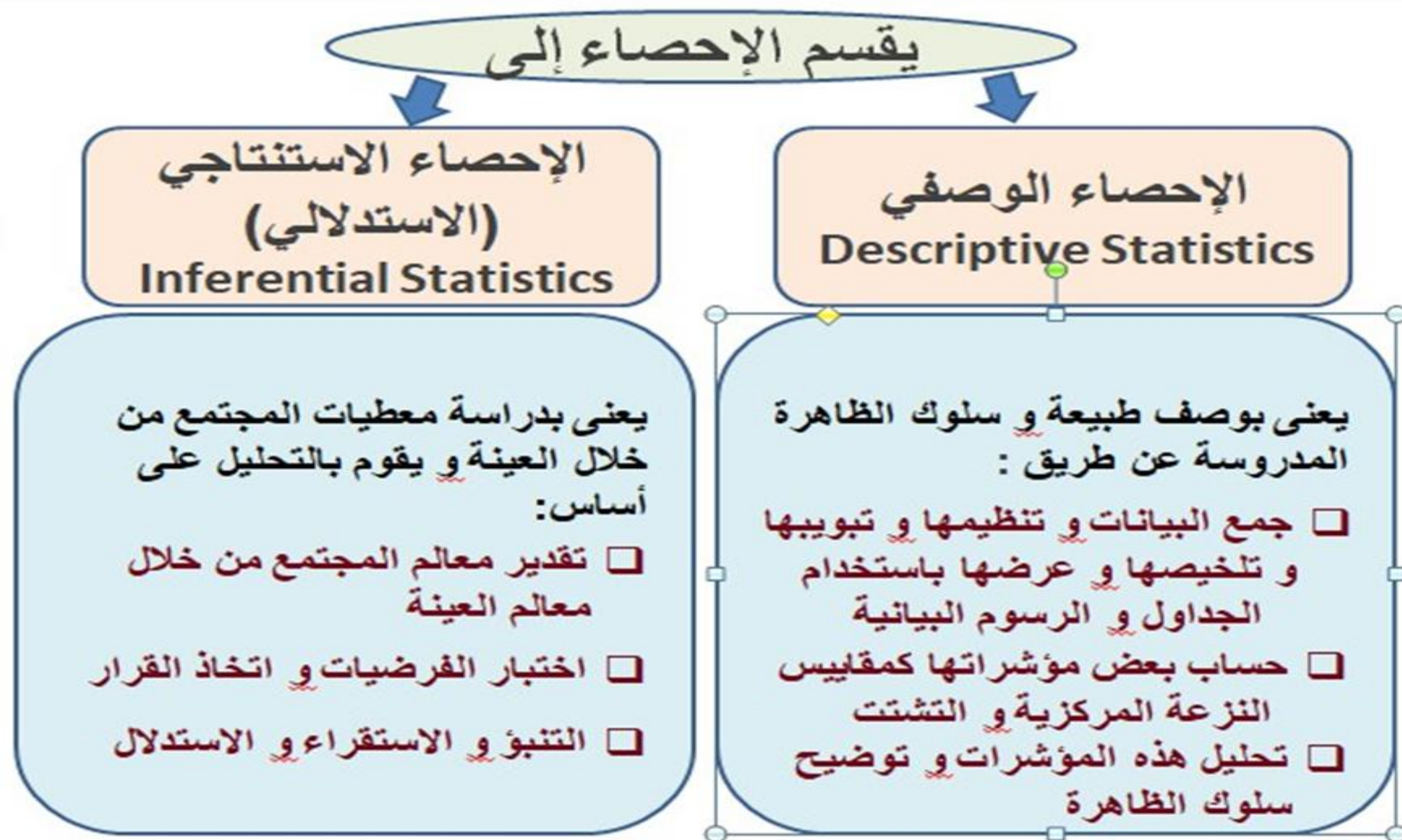
- الإحصاء علم: لأنه جزء من الرياضيات التطبيقية يعالج سلوك الأرقام ويستخلص القوانين الإحصائية.
- الإحصاء فن : بعدّه أسلوب وطريقة علمية لأصول جمع الحقائق بالأرقام ومعالجتها بمنهجية علمية.

ينقسم علم الإحصاء إلى نوعين:

الإحصاء الوصفي

الإحصاء الإستدلالي

كليات الخليج
Gulf Colleges



وظائف علم الإحصاء

1. عرض المعطيات و المشاهدات عن الظاهرة المدروسة بشكل واضح
2. تلخيص هذه المشاهدات و المطيات على شكل قيم ذات معنى (متوسطات ، نسب ، معدلات ، ... و غيرها)
3. وضع أسس للمقارنة بين متغيرات الظاهرة المدروسة
4. يساعد في وضع الفرضيات البحثية و اختبارها و تطوير النظريات العلمية
5. يساعد في معرفة اتجاه تطور الظاهرة المدروسة في المستقبل
6. يساعد متخذي القرار في وضع البرامج والخطط عن طريق ما يوفره من بيانات

أنواع المتغيرات وسلالم القياس

المتغيرات العشوائية

يتكون المجتمع (الجمهرة) الإحصائي من مجموعة من الأفراد أو العناصر (مجتمع غير بشري) تشترك فيما بينها ببعض الخصائص أو الميزات (سكان مدينة – طلاب كلية – مجموعة مرضى – إنتاج مصنع)

عند دراسة هذا المجتمع يجب أن نحدد الهدف من الدراسة (حساب متوسط الدخل – قياس مستوى السكر في الدم – متوسط تكاليف العمل الجراحي)

إذن... المتغير العشوائي – هو الوسيلة الرياضية التي نعبر بها عن المؤشر المدروس (الدخل – مستوى السكر - التكاليف) لعناصر المجتمع الإحصائي و التي يجب دراستها و تحليلها

كليات الخليج
Gulf Colleges

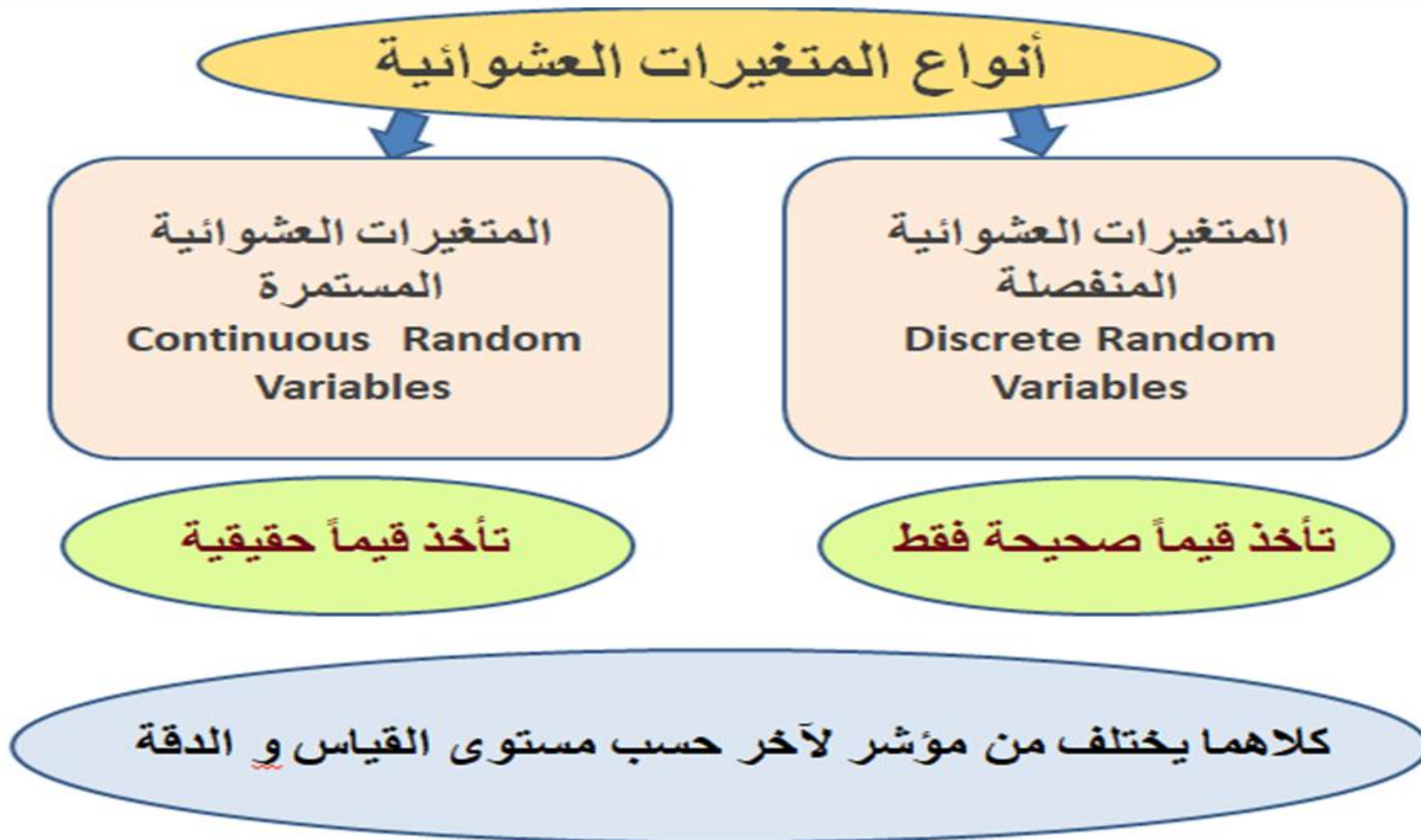
صفات أو خواص المتغير العشوائي

1. العشوائية : تتغير القيم من عنصر لآخر بشكل عشوائي

2. لكل عنصر من عناصر المجتمع (العينة) قيمة وحيدة

3. يعرف على كافة الأفراد و يأخذ قيم حقيقية أي في (المجموعة R)

4. لكل متغير عشوائي مدى محدد من القيم (فمثلاً مستوى ضغط الدم [120 – 80] ملم زئبقي



كليات الخليج
Gulf Colleges

كليات الخليج
Gulf Colleges

1- المتغير Variable :- يقصد به اي صفة او عنصر قابل للتغير في النوع والكم من فرد الى أخر في نفس المجتمع ويكون المتغير اما :-

A-متغيرات وصفية او نوعية Qualitative Variable

وهي الصفة التي لا يمكن قياسها مباشرة بأرقام عددية لان الفرق بين المفردات تكون في النوع وليس في الكم ومن الامثلة على ذلك (الصحة ، اللون ، الذكاء ، والجنس ، والحالة الاجتماعية)

B- صفة كمية Quantitative Variable : وهي الصفة التي يمكن قياسها مباشرة بأرقام عددية كالاختلاف بين الافراد في الطول والوزن ومستوى الهيموكلوبين والهرمونات وعدد خلايا الدم الحمراء ومستوى الدهون في مصل الدم Lipid profile (TC ، TG ،HDL-C ، LDL-C ، VLDL) ويمكن قياسها بوحدات القياس المختلفة كالسنتيمتر والكيلوغرام (mg , pg ,g) وتنقسم المتغيرات الكمية الى :-

1- متغيرات متصلة او مستمرة Continuous variable

المتغير المتصل هو المتغير الذي تأخذ كل مفردة قيمة رقمية او كسر بين حدي المتغير الكلي فلو فرضنا اطوال الطلبة يتراوح بين (130.5 و 170 سم) ،كمية الهيموكلوبين (12.5 – 14 ملغم لكل لتر من الدم)

2- متغيرات غير متصلة او مستمرة Discontinuous Variable

هي المتغيرات التي تأخذ المشاهدة او المفردة فيها قيم متباعدة او متقطعة غير مستمرة اي هو الذي لا تأخذ كل مفردة فيه قيمة كسرية بل لا تزيد قيمة المتغير او تنقص بأقل من واحد فعدد الطلاب عدد الكتب كلها متغيرات غير متصلة او مستمرة

المشاهدة Observation.

تعتبر المشاهدة ك بمثابة المواد الاولية التي يتعامل معها الباحث فأذا اراد باحث ان يقيس مستوى الكلوكوز في مصل دم احد الجرذان ولنفرض ان مستوى الكلوكوز في مصل دم هذا الجرذ هو (120 ملغم/ 1مل) فأن هذا العدد يمثل المشاهدة ، لذا فأن المشاهة هي سجل رقمي لحادثة وان مجموع المشاهدات تكون البيانات Data .

المجتمع Population .

المجتمع من الناحية الاحصائية يمثل جميع الافراد او العناصر التي تشترك في صفة متغير واحد او اكثر تميزه تماماً عن بقية المجتمعات ويتعلق مفهوم المجتمع بالهدف المحدد للبحث الاحصائي فقد يشكل طلبة جامعة المنارة مجتمعاً ،

والمجتمع هو عبارة عن جميع القيم التي يمكن ان يأخذها المتغير ، فمثلاً عند دراسة مستوى الهيموكلوبين في دم طلبة جامعة المنارة وصفة مستوى الهيموكلوبين في دم طلبة جامعةالمنارة هي متغير تأخذ مدى معين لمجتمع طلبة جامعةالمنارة

، والمجتمع اما ان يكون

A-مجتمع محدود Finite Population

وهو المجتمع الذي يمكن حصر مفرداته كما هو الحال في مستوى الهيموكلوبين في دم طلبة جامعة المنارة او عدد ردهات المرضى في مستشفى تشرين .

B-مجتمع غير محدود Infinite Population

هو المجتمع الذي من الصعب او المستحيل حصر مفرداته مثل عدد البكتريا في مستعمرة بكترية او حقل معين .

كليات الخليج
Gulf Colleges

c-الثوابت: فهي السمات والخواص التي لا تتغير، وهي تصف ماهية المواد في ظروف معينة،

d-معلمة المجتمع population parameter: هي الثابت الذي يصف المجتمع، وهو عبارة عن مقياس سمة، مثل: معدل المجتمع.

العينة Sample :-

العينة هي جزء المجتمع وهي عبارة عن مجموعة من المشاهدات اختير بطريقة ما من المجتمع حيث ان دراسة المجتمع ككل قد يكون صعباً ويحتاج الى وقت وجهد ومال لذا فقد استعيض عن دراسة المجتمع بدراسة العينة ومنها نستطيع ان نستنتج خواص المجتمع الذي اخذت منه العينة ، فقد تكون العينة انسان او حيوان او نبات او جزء معلوم من نبات معين تجري عليه التجارب في المختبرات والعينة هي احدى ادوات البحث العلمي .

e-الإحصاءات statistic: هو متغيّر تتحدّد قيمته من العيّنة أي أنه مقياس سمة تجد قيمته لأفراد العيّنة فقط، مثال: عمر المصباح.

كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة الثانية

أنواع سلالم القياس(المقاييس)

المقاييس - هي التي تحدد طريقة تلخيص و وصف الظواهر بناءً على قواعد محددة و هذه القواعد:

- القواعد الأساسية لاستخدام الأرقام (كمية – تمييزية)
- الخواص الرياضية للمقياس الناتج من الأرقام
- العمليات الإحصائية الممكن استخدامها مع المقياس

بناءً عليه هناك أربع أنواع للمقاييس (المستوى)

- المقياس الاسمي
- المقياس الرتبي
- المقياس الفتروي
- المقياس النسبي

كليات الخليج
Gulf Colleges

المقياس الاسمي Nominal Scale

- هو أدنى مستويات القياس
- تستخدم فيه الأعداد للتمييز بين الأشياء
  (تجميع الأشياء التي تشترك بخاصية معينة واحدة أو أكثر)
- ليس للعمليات الحسابية على الأرقام أي معنى
  (لأن الأرقام هنا لا تعبر عن كمية ما يحتويه الشيء من الظاهرة)
- المجموعات المتشابهة لا تتميز عن غيرها بالأهمية أو الترتيب

توزيع مجموعة من الأفراد حسب النوع و الإصابة بمرض معين

- (1) ذكر مريض
- (2) ذكر سليم
- (3) أنثى مريضة
- (4) أنثى سليمة

الأرقام هنا للتصنيف و لا تعبر عن كميات معينة

المقياس الرتبي Ordinal Scale

- يأتي في المرتبة الثانية بعد المستوى الاسمي
- يسمح بترتيب الصفات المدروسة بغض النظر عن تساوي الفروق بينها
- الأرقام التي تعبر عن الرتب لا تعبر عن مقدار توفر الظاهرة في المشاهدة
- لا يمكن إجراء أي عملية حسابية على الرتب و لكن يمكن حساب عدد التكرارات و حساب بعض المؤشرات (وسيط – منوال ....)

ترتيب مجموعة من الطلاب حسب تسلسل درجاتهم تنازلياً

- (1) الأول
- (2) الثاني
- (3) الثالث
- (4) الرابع ....

الفروق بين الترتيبات لا تساوي الفروق بين الدرجات

كليات الخليج
Gulf Colleges

المقياس النسبي Ratio Scale

- يتمتع بجميع الصفات السابقة
- الصفر قيمة عددية حقيقية مطلقة و يعني انعدام الظاهرة
- يمكن إجراء جميع العمليات الحسابية
- يمكن حساب جميع المؤشرات الإحصائية
- يعتبر أعلى مستويات القياس

ترتيب مجموعة من المرضى حسب مستوى السكر

بناءً على ما ذكر أعلاه (أنواع المتغيرات و مستوى القياس) يمكن أن نحدد فيما إذا كنا سنستخدم الطرائق المعلمية أو اللامعلمية في الدراسة و التحليل

المقياس الفتروي Interval Scale

- أدق من سابقيه و يتمتع بنفس خصائصهما
- يتمتع بواحدات قياس متساوية تعبر عن قيمة الظاهرة و يمكن جمعها أو طرحها
- لا يمكن إجراء عملية القسمة على القياسات لعدم وجود الصفر المطلق (الصفر قيمة نسبية و ليست مطلقة)
- يمكن حساب جميع المؤشرات (متوسطات ، معاملات ارتباط و غيرها)

ترتيب مجموعة من العمال حسب شرائح الدخل

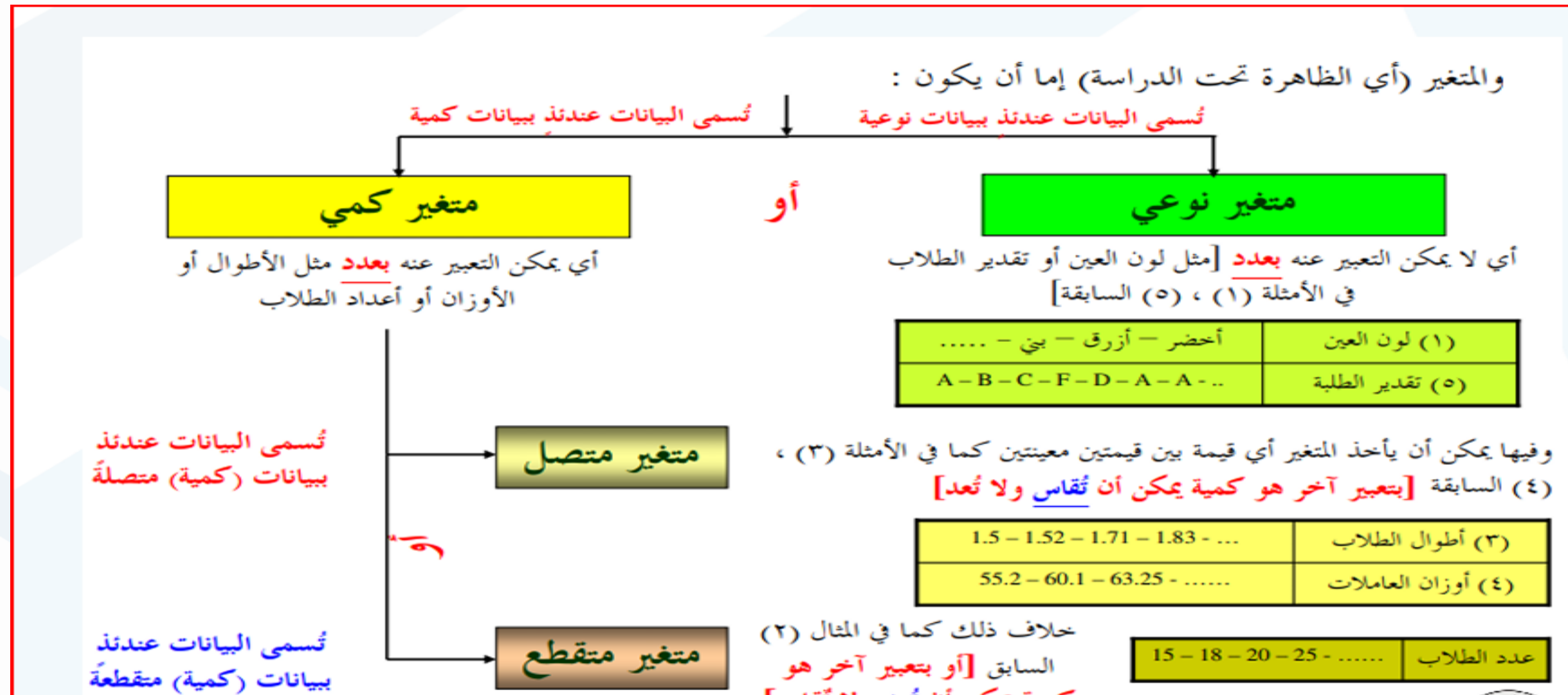
- (0) الشريحة الأولى [10-15]
- (1) الشريحة الثانية [15-20]
- (2) الشريحة الثالثة [20-25]
- (3) الشريحة الرابعة [25-30]

الفروق بين الدرجات تساوي الفروق بين الشرائح و الصفر ليس قيمة مطلقة

(٤) البيانات

يمكن ببساطة تعريف البيانات على أنها مجموعة من " المشاهدات أو القياسات " التي تخص الظاهرة تحت الدراسة ، والكمية التي نقوم بمشاهدتها أو قياسها تُسمى بالمتغير وعادةً نرمز له برمز مثل .. x , y , A , B ، فمثلاً :

مثال	العملية الإحصائية : دراسة	البيانات (القياسات أو المشاهدات)	المتغير x
(١)	لون العين لبعض الأطفال حديثي الولادة	أخضر – أزرق – بني –	لون العين
(٢)	عدد الطلاب في فصول مدرسة	15 – 18 – 20 – 25 – 17 -	عدد الطلاب
(٣)	أطوال مجموعة من الطلاب في فصل ما (بالمتر)	1.5 – 1.52 – 1.71 – 1.83 -	طول الطالب
(٤)	أوزان بعض العاملات بمصنع معين (بالكيلوجرام)	55.2 – 60.1 – 63.35 – 70.52 -	وزن العاملة
(٥)	تقديرات عدد من الطلاب في مقرر الإحصاء	A – B – C – D – F – A – C – B -	تقدير الطالب



والمتغير (أي الظاهرة تحت الدراسة) إما أن يكون :

تُسمى البيانات عندئذ ببيانات نوعية

تُسمى البيانات عندئذ ببيانات كمية

متغير نوعي

أو

متغير كمي

أي لا يمكن التعبير عنه **بعدد** [مثل لون العين أو تقدير الطلاب في الأمثلة (١) ، (٥) السابقة]

(١) لون العين	أخضر – أزرق – بني – .....
(٥) تقدير الطلبة	A – B – C – F – D – A – A - ..

أي يمكن التعبير عنه **بعدد** مثل الأطوال أو الأوزان أو أعداد الطلاب

متغير متصل

تُسمى البيانات عندئذ ببيانات (كمية) متصلةً

وفيها يمكن أن يأخذ المتغير أي قيمة بين قيمتين معينتين كما في الأمثلة (٣) ، (٤) السابقة [بتعبير آخر هو كمية يمكن أن تُقاس ولا تُعد]

(٣) أطوال الطلاب	1.5 – 1.52 – 1.71 – 1.83 - …
(٤) أوزان العاملات	55.2 – 60.1 – 63.25 - ……

أو

متغير متقطع

تُسمى البيانات عندئذ بيانات (كمية) متقطعةً

خلاف ذلك كما في المثال (٢) السابق [أو بتعبير آخر هو

عدد الطلاب	15 – 18 – 20 – 25 - ……

كليات الخليج
Gulf Colleges

ذكرنا في الباب السابق (الباب الأول) ما هي البيانات [هي مجموعة المشاهدات أو القياسات التي تخص ظاهرة معينة تحت الدراسة] وعرفنا المتغير على أنه تلك الكمية التي نقوم بمشاهدتها أو قياسها ، كما ذكرنا أن البيانات إما أن تكون : نوعية أو كمية ، حيث :

(أ) البيانات النوعية : هي تلك البيانات التي لا يمكن التعبير عن متغيرها بعدد (أي بيانات غير رقمية) ، مثل :

- لون (أو نوع) السيارات الموجودة في موقف ما [أحمر – أبيض – أسود – ....]
- الحالة الاجتماعية للسيدات في محافظة معينة [متزوجة – عزباء – مطلقة – أرملة – منفصلة]
- رأيك في قرار خاص بالمؤسسة التي تعمل بها [أوافق بشدة – أوافق – أعترض – أتحفظ – ...]
- وغيره من مثل هذه الأمثلة .

(ب) البيانات الكمية : هي تلك البيانات التي يُعبر فيها عن المتغير بعدد (أي بيانات رقمية) ، وهذه البيانات بدورها تنقسم إلى :

(ب) البيانات الكمية : هي تلك البيانات التي يُعبر فيها عن المتغير بعدد (أي بيانات رقمية) ، وهذه البيانات بدورها تنقسم إلى :

(ب – ١) بيانات كمية متصلة : وفيها يمكن أن يأخذ المتغير أي قيمة بين قيمتين (أي بيانات يمكن أن تُقاس ولا تُعد ، مثل :

- أطوال الطلاب في إحدى المدارس .
- أوزان العاملات بإحدى المصانع .
- الدخل السنوي لمنسوبي مؤسسة معينة .
- وغيره من مثل هذه الأمثلة .

(ب – ٢) بيانات كمية متقطعة : وفيها يمكن أن يأخذ المتغير قيمة على (إما ..... أو ...... وليس أي قيمة بينهما) ، وبتعبير آخر هي بيانات يمكن أن تُعد ولا تُقاس ، مثل عدد طلاب الفصول المختلفة في مدرسة ما

والبيانات المنفصلة إما أن تكون نوعية أو كمية متقطعة

تمارين محلولة الفصل الاول

التمرين 01: – ما المقصود بكل من: الإحصاء الوصفي والإحصاء الإستدلالية وما الفرق بينهما؟

– ما المقصود بأسلوب الحصر الشامل وأسلوب المعاينة؟ وما هي الاعتبارات التي تحكم المفاضلة بين الأسلوبين؟

– ما الفرق بين المجتمع الإحصائي المحدود وغير المحدود؟

– ما الفرق بين المتغير الكفي و المتغير الكمي؟

– لماذا يفضل الباحث البيانات الأولية عن البيانات الثانوية؟

– وما الفرق بين أسلوب جمع البيانات عن طريق المقابلة و عن طريق الهاتف؟

التمرين 02: بين أي الأسلوبين (الحصر الشامل أو العينة) يستخدم في الحالات التالية:

- عند التعرف على جودة المنتجات في الخطوط الإنتاجية بأحد المصانع؛
- عند إجراء تحاليل دم أحد المرضى؛
- التعداد السكاني؛
- عند فرز أصوات المنتخبين يوم إجراء الانتخابات الرئاسية؛
- عند معرفة التعداد الزراعي؛
- عند معرفة مدى فعالية دواء ما، لمعالجة مرض معين؛
- عند إحصاء عدد المتزوجين والطلاق، والأرامل والأيتام؛
- عند التعرف على آراء المستهلكين نحو منتوج جديد تم طرحه من طرف مؤسسة ما في السوق؛

حل التمرين 02: تعيين أي الأسلوبين يستخدم الحصر الشامل أو العينة:

الحالات	أسلوب الحصر الشامل	أسلوب العينة
تعرف على جودة المنتجات	________	عن طريق العينة
تحليل دم أحد المرضى	________	عن طريق العينة
التعداد السكاني	أسلوب الحصر الشامل	________
فرز أصوات المنتخبين	أسلوب الحصر الشامل	________
التعداد الزراعي	أسلوب الحصر الشامل	________
مدى فعالية دواء لمرض معين	________	عن طريق العينة
عند إحصاء عدد المتزوجين والطلاق والأرامل والأيتام	أسلوب الحصر الشامل	________
معرفة آراء المستهلكين نحو منتوج جديد	________	عن طريق العينة

تمرين 3:

صنف السمات والقيم الاتية الى متغيرات و ثوابت - :دخل الفرد العامل اليومي؛ العدد الذري للهيدروجين؛ كمية المحصول الزراعي؛ درجات الحرارة، قوة جاذبية الأرض، سرعة الضوء، عدد أيام الأسبوع، عدد دقائق الساعة، سرعة السيارة، عدد دول العالم ، طول المسافة بني مدينة اللاذقية ودمشق عدد الوالدات في مستشفى التوليد ، عدد حوادث المرور في اللاذقية ، عدد محركات السيارة، فصيلة الدم لدى شخص مريض؛

المتغيرات	الثوابت
-دخل الفرد العامل اليومي؛ كمية المحصول الزراعي؛ درجات الحرارة، سرعة السيارة، عدد الوالدات مستشفى التوليد عدد حوادث المرور في اللاذقية	العدد الذري للهيدروجين؛ قوة جاذبية الأرض، سرعة الضوء، عدد أيام الأسبوع، عدد دقائق الساعة عدد دول العالم ، طول المسافة بني مدينة اللاذقية ومدينة دمشق ، عدد محركات السيارة، فصيلة الدم لدى شخص مريض.

تمرين 4: حدد نوع الخاصية المدروسة ونوعية المتغير عندما يتعلق الدراسات التالية :

-الجنس، الدخل السنوي، ترتيب اللاعب في لعبة الغولف ، عدد السيارات المباعة في أسبوع ما، ترتيب المبيعات السنوية لإحدى الشركات الصناعية، لون العيون؛ قيمة الاتصالات الهاتفية اليومية، عدد المكالمات الهاتفية، درجات الحرارة، ترتيب الناجحين في مسابقة الدكتوراه؛ مكان الإقامة.

حل التمرين 04: تحديد الخاصية المدروسة ونوعية المتغير:

الخاصية المدروسة	نوعية المتغير	الخاصية المدروسة	نوعية المتغير
الجنس ، لون العيون، مكان الإقامة	**كيفي إسمي**	الدخل السنوي، قيمة الاتصالات الهاتفية، درجات الحرارة،	**كمي مستمر**
ترتيب اللاعب، ترتيب الناجحين في مسابقة الدكتوراه، ترتيب المبيعات	**كيفي ترتيبي**	عدد السيارات المباعة، عدد المكالمات الهاتفية	**كمي منفصل**

تمرين 5: تمت تنظيم مسابقة الدكتوراه جامعة تشرين 2، بكلية الاقتصاد ، حيث تم استقبال 500 طالب، وتم وضع مجموعة من الشروط وهي كالتالي

1- :ان يكون تخصص الطالب ضمن التخصصات التالية: التسويق، المحاسبة ، احصاء ، إدارة الأعمال، مالية؛

2-أن يكون تقدير الطالب يف شهادة الماستر جيد فما فوق؛

3-أن الا يتجاوز سن الطالب يوم المتحان 35 سنة؛

4-عدم دخول الطالب أكثر من خمس مقررات للدورات التكميلية خلال مساره الجامعي؛

MANARA UNIVERSITY

5-أن يكون معدل الطالب خلال مرحلة لليسانس يساوي أو يفوق 12 من عشرين

؛ المطلوب: حدد المجتمع الإحصائي و الوحدة الإحصائية، حدد الصفات المدروسة ونوعيتها؟

رقم الجواب	الصفة المدروسة	نوع المتغير
1	التخصص	كيفي
2	التقدير	كيفي
3	عمر الطالب	كمي مستمر
4	عدد الدورات التكميلية	كمي منفصل
5	معدل الطالب	كمي مستمر

نمرين 6: حدد كل من: اسم المتغير - نوع البيانات – مجتمع الدراسة – حجم العينة

1- : في دراسة على عينة من 100 طفل دون سن السادسة لمعرفة اللعبة التي يفضلها كل منهم؛

2-في عينه من 500 مريض السرطان يراد قياس مدة الحياة بعد التشخيص

3- في دراسة لتحديد فصيلة الدم لعينة من 200 مريض بالمستشفى؛

4-في دراسة على 400 مريض لتحديد عدد الساعات اليت قضوها داخل المستشفى حيت تماثلوا للشفاء من مرض كوفيد 19.

5- في دراسة لتحديد أوزان عينة من 45 طفل حديث الوالدة بالكيلو جرام يف مستشفى لتوليد خلال الأسبوع الأول من شهر كانون اول

6-في دراسة لقياس مقدار النقص يف الوزن بالكيلو جرام لعينة من 30 لاعب في فريق الملاكمة ذو الأوزان الكبيرة و يتبعون نظام غذائي معين؛

7-في دراسة على عينة من 100 النساء الحوامل لقياس الوقت بالساعات الذي تقضيه في المستشفى قبل الوالدة؛

8-في دراسة لتحديد عدد المرات التي أصيب فيها الطفل في سن 5 سنوات بالتهاب الأذن على عينه من 50 طفل يف مدينة اللاذقية

حل التمرين 6

الرقم	اسم المتغير	نوع البيانات	مجتمع الدراسة	حجم العينة
01	اللعبة	كيفي	الأطفال	100 طفل
02	مدة الحياة	كمي مستمر	المرضى بالسرطان	500 مريض
03	فصيلة الدم	كيفي	المرضى بالمستشفى	200 مريض
04	عدد الساعات	كمي منفصل	المرضى	400 مريض
05	أوزان-الوزن	كمي مستمر	الأطفال	45 طفل
06	الوزن	كمي مستمر	اللاعبين	30 لاعب
07	الحمل	كمي منفصل	النساء الحوامل	100 من النساء الحوامل
08	عدد مرات الإصابة	كمي منفصل	الأطفال	50 طفل

كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة الثالثة

البحث الإحصائي:

يعرّف بأنه استقصاء منظم ومضبوط وتجريبي وناقد للافتراضات حول طبيعة العلاقات بين المتغيّرات لظاهرة ما بهدف اكتشاف المعرفة بطريقة منظمة لجمع الحقائق وتنسيقها بأسلوب علمي ومنطقي وصولاً للقرار السليم وتعميم النتائج الحاصلة على المجتمع الإحصائي.

كليات الخليج
Gulf Colleges

خطوات البحث الإحصائي:

1. اختيار الموضوع
2. تحديد مشكلة البحث
3. وضع الفرضيات "فروض البحث"
4. تصميم البحث "خطة البحث"
5. استمارة البحث
6. مصادر البيانات
7. أسلوب جمع البيانات
8. ترتيب وتبويب البيانات
9. تحليل البيانات الإحصائية
10. القرار الإحصائي:

كليات الخليج
Gulf Colleges

1- اختيار الموضوع:

. وفيما يلي أهم الاعتبارات الواجب مراعاتها عند اختيار موضوع البحث:

1- الاهتمام الشخصي في اختيار الموضوع بهدف الاستمرار في بحث معين والتركيز على ناحية معينة يريد الباحث التركيز مثلاً على العلاقة بين هيكل الشركة وبين الوجهة المهنية فيها.
2- التركيز على بعض المشكلات الاجتماعية والتربوية والنفسية مثلاً دراسة التلوث وتأثيراته وحماية البيئة.
3- الاهتمام على بعض الجوانب النظرية لبعض المشكلات بهدف معرفة مدى ملاءمتها وتأثيراتها ومدى فائدتها في إيجاد بعض الحلول مثلاً دراسة أثر البيروقراطية والإدارة ونظريات السلوك الإنساني.

-2: تحديد مشكلة البحث:

مشكلة البحث: عبارة عن شكوك لها ما يبررها من أسباب ودوافع تجعل الظاهرة في وضع مخالف ومغاير لنمطيتها الطبيعية وما يجب أن تكون عليه.

كليات الخليج
Gulf Colleges

وتتوقف تحديد مشكلة البحث على عوامل عدة منها:

1- وضوح الرؤية بالنسبة لحقائق المشكلة ومتغيّراتها.
2- موضوعية الدراسة وإمكانية تنفيذها.
3- إمكانية الحصول على البيانات بشكل دقيق وبأقل التكاليف.
4- توفر الإمكانيات المادية.
5- التوقيت اللازم لإنجازها.
6- طبيعة الظاهرة المدروسة.
7- الهدف من البحث.
8- أنواع المتغيّرات المراد دراستها / كمية / نوعية ... الخ.
9- التخصص والميل العلمي ونوعيته.

-3: وضع الفرضيات "فروض البحث" :

مشكلة البحث هي عبارة عن تساؤلات فإن الفرضية هي حلول مقترحة وبدائل وتفسيرات مؤقتة وتخمين واستنتاج يتطلب إثباته بالطريقة العلمية، أي بمعنى آخر الفرضية هي بدائل حلول مقترحة كما تعرف "تفسير مؤقت أو محتمل للعوامل والأحداث والظروف التي يسعى الباحث إلى تفسيرها".

كليات الخليج
Gulf Colleges

هذا وتأخذ صياغة الفرضية عدة أشكال أهمها:

1- طريقة الإثبات الطردية:
2- طريقة الإثبات العكسية:
3- طريقة صياغة النفي: " ؟

وفيما يلي أهم الاعتبارات الواجب مراعاتها في فروض البحث:

1- يجب أن تكون الفرضية دقيقة في صياغتها وواضحة في معناها بحيث لا تؤدي المفردات إلى ازدواجية في المعنى أو أن تكون قابلة للتأويل.
2- أن تكون منطقية وموضوعية وقابلة للاختبار والتجريب وأيضاً قابلة للتحقق والإثبات.
3- يجب أن تصاغ فروض البحث على شكل جمل استفهامية.
4- يجب أن تربط فروض البحث بين متغيّرين أو أكثر.
5- يجب ان تتضمن فروض البحث على مفهوم ضمني مؤداه إمكانية قياس المتغيّرات.
6- يجب أن تكون فروض البحث محدّدة بما ينسجم وطبيعة المتغيّرات المدروسة.

كليات الخليج
Gulf Colleges

-4: تصميم البحث "خطة البحث":

تصميم البحث عبارة عن مخطط تنظيمي لخطوات ومراحل البحث في شقيها النظري والإجرائي. يعرف على أنه دليل أو مجموعة من الخطط لجمع المعلومات والبحث ذو التصميم المثالي هو الذي يجمع أقصى قدر من المعلومات مع الإنفاق الأدنى للوقت والموارد.

-5: استمارة البحث:

الاستمارة بالتعريف: عبارة عن كشف يتضمن الأسئلة المراد تقديمها من قبل الباحث بهدف جمع البيانات والمعلومات حول الظاهرة المراد دراستها. وتعدّ من العوامل الجوهرية في إنجاح عملية البحث

وللاستمارة أنواع / للاستبيان أنواع أهمها:

1- **الاستبيان المغلق**: وتكون فيه الإجابات على الأسئلة محدّدة بعدد من الخيارات /نعم/ لا/ موافق/ غير موافق. ويتميز بالبساطة وسرعة تفريغ البيانات وقلة التكاليف. ومن عيوبه:

2- **الاستبيان المفتوح**: يعطي الفرد حرية للإجابة على الأسئلة الواردة في الاستبيان بدلاً من حصر إجابته في عدد محدّد من الخيارات.

4- **الاستبيان المغلق ـ المفتوح**: عبارة عن استبيان يتضمن خليطاً من الأسئلة الواردة في الاستبيان المغلق- المفتوح

أهم الاعتبارات الواجب مراعاتها عند تصميم الاستبيان.

1- وضوح وسهولة الأسئلة المقدّمة و لا تتطلب الرجوع بالذاكرة.
2- وضع أقل عدد ممكن من الأسئلة
3- الابتعاد عن الأسئلة الحرجة والحساسة.
4- عدم تقديم الأسئلة التي تحتاج إلى قيام الفرد بالعمليات الحسابية.
5- يفضل وضع الأسئلة التي تتطلب إجابات محدّدة ومختصرة.
6- ضرورة صياغة الأسئلة بلغة واضحة ومفهومة.
7- وجوب الإحاطة بسرية البيانات وعدم الإدلاء بها.
8- ضرورة تحديد واحدات القياس. وتحديد المفاهيم بشكل جيد.
9- ضرورة وضع تفسيرات وتوضيحات للمصطلحات

a- فوائد الاستبيان:

1- يعدّ وسيلة أقل كلفة من وسائل جمع المعلومات الأخرى.
2- يتطلب مهارة أقل لإنجازه من تلك التي تتطلبها المقابلة مثلاً.
3- إمكانية توزيعه على أكبر عدد من الإفراد في الوقت نفسه.
4- إعطاء ثقة للفرد في الإجابة على الأسئلة دون خوف وكذلك منحه الوقت الكافي..

b - سلبيات الاستبيان:

1- انخفاض نسبة الحصول على بيانات عند إرساله بالبريد أو بوسائل أخرى.
2- عدم إمكانية استخدامه في المجتمعات التي ما تزال نسبة الأمية فيها مرتفعة .
3- احتمال الإجابة بمعلومات خاطئة على الأسئلة الواردة في الاستبيان نظراً لعدم فهم الفرد مضمون الأسئلة المقدّمة.

Gulf Colleges كليات الخليج

6: مصادر البيانات :

1-**المصدر غير المباشر:** عبارة عن البيانات التي يمكن الحصول عليها دون قيام الباحث بالقياس المباشر، أي يعتمد على البيانات المتوافرة في السجلات والدوريات والمجلات والكتب والمخطوطات

2-**المصدر المباشر:** وهنا يقوم الباحث بنفسه بالحصول على البيانات اللازمة من خلال إجراء القياسات أو التجارب المخبرية مثال: الإنتاجية – القوة الشرائية وعلاقتها بالأسعار... الخ.

كليات الخليج
Gulf Colleges

7: أسلوب جمع البيانات:

1- أسلوب الحصر الشامل:

ينطوي هذا الأسلوب في الدراسة على أخذ جميع مفردات المجتمع الإحصائي دون تجاهل أي مفردة، ويعتمد هذا الأسلوب في مجال التعدادات السكانية والمساكن والحيازات الزراعية والصناعية

2 - أسلوب العيّنات:

العيّنة جزء من المجتمع الأصلي المراد دراسته وتحديد خصائصه وهي تمثل نسبة مئوية من حجم المجتمع الإحصائي، أي تسمح لنا بدراسة الكل عن طريق الجزء.

وللعينات العديد من الفوائد أهمها:

وللعينات العديد من الفوائد أهمها:

1-توفر الوقت والجهد والتكاليف.
2-تفيد في دراسة بعض الظواهر التي لا تشترط الحصول على معلومات دقيقة.
3-يعدّ الأسلوب الوحيد لدراسة بعض الظواهر مثال / تعداد كريات الدم/.
4-تفيد في التأكد من بعض الأخطاء الواردة في أسلوب الحصر الشامل.

: أسلوب جمع البيانات:

1- أسلوب الحصر الشامل:

ينطوي هذا الأسلوب في الدراسة على أخذ جميع مفردات المجتمع الإحصائي دون تجاهل أي مفردة، ويعتمد هذا الأسلوب في مجال التعدادات السكانية والمساكن والحيازات الزراعية والصناعية

2 - أسلوب العيّنات:

العيّنة جزء من المجتمع الأصلي المراد دراسته وتحديد خصائصه وهي تمثل نسبة مئوية من حجم المجتمع الإحصائي، أي تسمح لنا بدراسة الكل عن طريق الجزء.

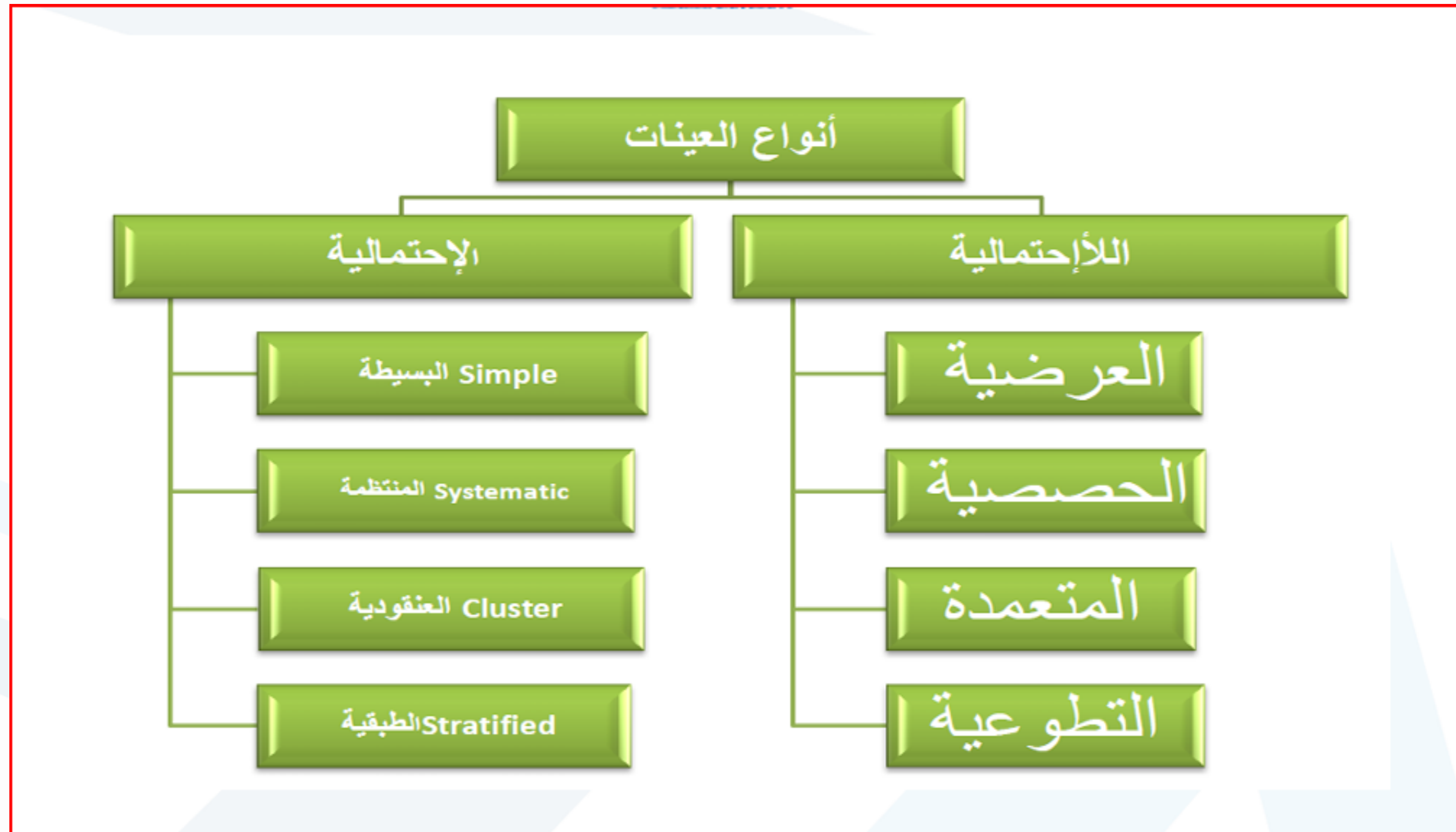
وللعينات العديد من الفوائد أهمها:

5- توفر الوقت والجهد والتكاليف.

6- تفيد في دراسة بعض الظواهر التي لا تشترط الحصول على معلومات دقيقة.

7- يعدّ الأسلوب الوحيد لدراسة بعض الظواهر مثال / تعداد كريات الدم/.

8- تفيد في التأكد من بعض الأخطاء الواردة في أسلوب الحصر الشامل.


أنواع العينات
الإحتمالية
اللاإحتمالية
البسيطة Simple
المنتظمة Systematic
العنقودية Cluster
الطبقيةStratified
العرضية
الحصصية
المتعمدة
التطوعية

4-1 انواع العينات (samples types)

ان العيّنة تعكس المعلوماتِ المجمعة عن بَعْض عناصر ِالمجتمع كما ان العيّنة الجيدة يَجِبُ أَنْ تعكس كُلّ الخصائص والمميزات للمجتمع وهذه العينات بالإمكان أَنْ تعكس الخصائص المهمة للمجتمع الذي سحبت منه بدرجاتِ دقّةِ مختِلفة . وأي عيّنة التي تعكس خصائص المجتمع بدقة تُدْعَى عيّنة تمثيلية (*representative*) ، كما ان أي عيّنة التي لَيستْ ممثلة لخصائصِ المجتمع تُدْعَى عينة متحيّزة (*biased*) وان مدى الثقة أَو الدقة بالإستنتاجاتِ المستخلصة للمجتمع يَعتمدُ على اسلوب اختيار العيّنةَ بشكل صحيح لِكي تُمثّلَ المجتمع ،

كليات الخليج
Gulf Colleges

وحيث ان هناك العديد مِنْ الطرقِ المتوفرةِ لسحب العينات وان الإختيار والمفاضلة ما بين طرقِ سحب العينات يَعتمدُ على مجموعة من العوامل منها .

1- طبيعة المشكلةِ أَو التحقيق .
2- مدى توفر إطاراتِ العينات الجيدةِ (قائمة كُلّ الأعضاء(العناصر) في المجتمع).
3- الميزانية أَو مدى توفر الموارد المالية .
4- المستوى المطلوب للدقةِ .
5- الطريقة التي سيتم فيها جمع البيانات مثل الإستفتاءاتِ أَو المقابلاتِ .

أخذ العينات sampling procedures

العينة السليمة هي العينة الممثلة للمجتمع الذي اختيرت منه، وعملية أخذ العينات بطريقة غير مناسبة واحدة من أسوأ الأخطاء الممكن ارتكابها. ويمكن تقسيم العينات وفقا لأسلوب اختيارها إلى نوعين هما العينات الاحتمالية والعينات غير الاحتمالية.

العينات الاحتمالية probability sampling: هي العينات التي يتم اختيار عناصرها وفقا لقواعد الاحتمالات، **بمعنى آخر هي التي يتم اختيار عناصرها من مجتمع الدراسة بطريقة عشوائية، بهدف تجنب التحيّز الناتج عن اختيار العناصر**، ومن أهم أنواع العينات الاحتمالية ما يلي:

1- العينة العشوائية البسيطة Simple Random Sample

وهي تلك العينة التي تسحب من مجتمع الدراسة بحيث يكون احتمال فرض ظهور اية مفردة من مفردات المجتمع الاحصائي في العينة متساوياً وبمعنى اخر تعني اعطاء كل فرد من المجتمع نفس الفرصة للظهور في العينة ويتم اختيارها كما يلي :

مثل استخدام طريقة البطاقات او القرعة

اذا كان لدينا (5) مرضى لاختبار دواء معين واردنا اختيار مريضين عشوائياً فما عدد الطرق الممكنة لاختيار مريضين لإجراء بعض الفحوصات .

ان عدد الطرق الممكنة $n\,Cr = \frac{n!}{r!(n-r)!}$

توافيق $5\,C2 = \frac{5\,!}{2!(5-2)!}$

1. **مثال العينة العشوائية البسيطة simple random sample**: تؤدي هذه الطريقة إلى احتمال اختيار أي عنصر من أفراد المجتمع كعنصر من عناصر العينة، كما أن لكل عنصر فرصة متساوية لاختياره ضمن العينة.

2. **العينة العشوائية المنتظمة systematic random sample المبدأ:** يتم فيها اختيار الوحدة الأولى(وتسمى بوحدة الابتداء) من العينة بطريقة عشوائية ثم يتم اختيار بقية الحالات على مسافات رقمية منتظمة أو متساوية بين الحالات(وتسمى بفترة السحب).

مثال : نريد سحب عينة بحجم n=50 حجمه مفردة من مجتمع N=500 وبالتالي تكون فترة السحب تساوي

$$k = \frac{500}{50} = 10$$

أي تسحب مفردة كل عشرة مفردات ويمكن تطبيق المتوالية العددية في سحب المفردات

$$k_n = a_1 + (n-1)r$$

أساس المتوالية(فترة السحب) الحد الأول (وحدة الابتداء)

المفردة الأولى = $k_1 = 5 + (1-1)10 = 5$

$k_{15} = 5 + (15 - 1)10 = 145$ المفردة رقم 15 =

وهكذا .... $k_{50} = 5 + (50 - 1)10 = 495$ المفردة رقم 50 في العينة=

من ميزاتها انها تسمح بمتابعة تغيرات نفس المفردة في الزمان والمكان خلال فترة او فترات زمنية متسلسلة

مثال العينة العشوائية المنتظمة :- وهي اختيار العينات بشكل منتظم من قائمة المجتمع حيث يتم اختيارها من خلال ترقيم عناصر المجتمع الاحصائي بحيث يتم تحديد قاعدة للاختيار تستند على تحديد اختيار العنصر الاول ولتبسيط الشرح لو كان مجتمع الاصل (100 مريض) وتريد اختيار (10 مرضى) لأجراء بعض الفحوصات عليهم فمثلاً تأخذ الارقام العشرة الاولى وتوضع في صندوق ويتم السحب ، فمثلاً حصلنا على الرقم (3) فيكون العينات العشرة المرضى هي كالآتي ،

$$5\,C2 = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1\ (3 \times 2 \times 1)} = 10$$ طرق مختلفة

بمعنى انه يوجد عشر بطاقات يكتب عليها اسم مريضين ويتم اختيار بطاقة من العشرة بطاقات عشوائياً .

فأذا كانت اسماء المرضى c,e,d,b,a فأن العشر بطاقات يكون مكتوب عليها de, ce ab , ac , ad , ae , bc , bd, be , cd, ويتم سحب اي بطاقة من العشرة .

المسافة= $+ \frac{100}{10} = 10 \frac{\text{مجموع المرضى}}{\text{مجموع العينة}}$

93, 83,73, 63, 53 , 43,33,23,13,3,,,

حيث ان الفاصلة فترة السحب = 10 بين مريض و آخر وتسمى هذه العينة المختارة (عينة منتظمة Systematic Sample) .

شرط المعانية : تجانس المجتمع المدروس.
المبدأ.اختيار عناصر العينة على مسافات متساوية
العشوائية تكمن في تحديد المفردة الاولى

كليات الخليج
Gulf Colleges

شكراً جزيلاً لكم

الاحصاء الحيوي

Biostatistics



المحاضرة الرابعة

كليات الخليج
Gulf Colleges

3. العينة العشوائية الطبقية stratified random sample: يتم في هذه الطريقة تقسيم المجتمع إلى عدة مجموعات نسمي كل منها طبقة تشترك بنفس الخواص (الجنس، العمر، ...).

المبدأ :يتم في هذه الطريقة تقسيم المجتمع غير المتجانس إلى عدة مجموعات (طبقات) تتسم كل منها بدرجة عالية من التجانس الداخلي

الاسلوب وتتم المعانية داخل كل طبقة بطريقة المعانية العشوائية

الشرط: عدم تجانس المجتمع المدروس

كليات الخليج
Gulf Colleges

الخطوات المتبعة لإخْتيارَ العينة الطبقية

1. تحديد عوامل السُحُب الطبقيّةِ مثل (جنس، عُمر، دخل، الخ.)
2. قسّمْ كامل المجتمع إلى الطبقاتِ وهي عبارة عن مجتمعات جزئية (sub populations) بالاستناد الى العوامل والمعايير الطبقيّةِ وبحجوم للطبقات قَدْ تَتفاوتُ فيما بينها .
3. إخترْ العددَ الضروريَ من الوحداتِ التي تَستعملُ كعينة عشوائيةَ بسيطةَ أَو عينات منتظّمةَ مِنْ كُلّ مجتمع جزئي وان العدد الضروري قَدْ يَعتمدُ على حجومِ المجتمع الجزئيةِ..

نسمي كل منها طبقة تشترك بنفس الخواص (الجنس، العمر،

عدد مفردات الطبقة x

عدد المفردات من كل طبقة =------------------------ * حجم العينة المطلوبة

عدد مفردات كل الطبقات في المجتمع

$$n_{starte} = \frac{n_k}{N} * n_m$$

كليات الخليج
Gulf Colleges

مثال : ترغب إدارة جامعة المنارة اختيار لجنة من طلاب الجامعة بحجم 50 طالب وطالبة اذا علمت ان عدد طلاب الجامعة للعام 2019-2020 1250 طالب وطالبة

حدد عدد الطلاب من كل كلية اذا ان عدد الطلاب في الكليات وفق الجدول الاتي:

الأداء	الهندسة	الإدارة	الصيدلة	الكلية
35	390	280	545	عدد الطلاب

545

عدد الطلاب من الصيدلة =- --------- *50= 22 طالب وطالبة وهكذا بالنسبة لباقي الكليات

1250

وفق الجدول الاتي

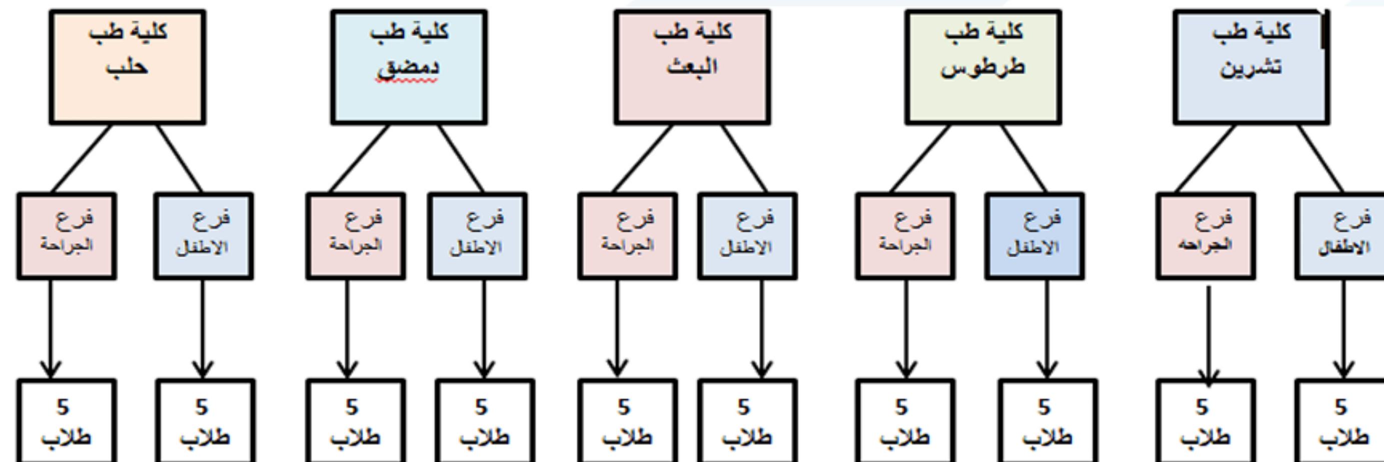
الكلية	عدد طلاب الكلية	عدد الطلاب من كل كلية
الصيدلة	545	22
إدارة الاعمال	280	11
الهندسة	390	16
الأداء	35	2
المجموع	1250	نتيجة تقريب الأرقام 51

4. العينة العشوائية العنقودية cluster random sample: يتم في هذه الطريقة تقسيم المجتمع إلى مقاطع، بعدها يتم اختيار بعض هذه المقاطع عشوائياً ومن ثم نختار جميع العناصر من تلك المقاطع.

كليات الخليج
Gulf Colleges

1- العينة العنقودية متعددة المراحل Multi-stage cluster sample

تعتبر المعاينة العنقودية احد الآليات التي يمكن استخدامها لاختيار العينات من خلال تقسيم المجتمع الى مجموعات او عناقيد على سبيل المثال نريد التعرف على مستوى التعليم الطبي في سورية فنختار 5 كليات طب ومن كل كلية فرعين ومن كل فرع 5 طلاب



كليات الخليج
Gulf Colleges

العينات غير الاحتمالية non probability sampling:

* **هي العينات التي يتم اختيار عناصرها بطريقة** غير عشوائية، حيث يقوم الباحث باختيار عناصر العينة بالصورة التي تحقق الهدف من المعاينة، مثل اختيار عينة من المزارع التي تنتج التمور من النوع السكري، وأهم أنواع العينات غير الاحتمالية:

1. **العينة العمدية purposive sample:** يتم اختيار الحالات بناء على هدف خاص لدى الباحث.
مثال تحليل محتوى مجلة محددة، دراسة متعمقة لبعض حالات التخلف العقلي.

2. **العينة الحصصية quota sample: تتطلب** معرفة مسبقة لمجتمع الدراسة من حيث تكوين المجموعات داخله، وعملية الاختيار في كل مجموعة لا ترتبط بقواعد معينة ولكن لقناعة الباحث بشرط أن تمثل كل مجموعة في العينة حسب تمثيلها في مجتمع الدراسة. مثال تحديد الباحث فئات المجتمع (ذكور وإناث) ثم يختار عدد ثابت من كل فئة إذ يقرر اختيار عشرة ذكور وخمس إناث

8- -ترتيب وتبويب البيانات:

هي عبارة عن تنظيم البيانات بعد تفريغها من الاستبيان في صورة جداول رقمية بحسب المتغيّرات المراد دراستها بصورة تسهل عملية التحليل الرياضي و الإحصائي وفق المنهجية المحدّدة مسبقاً وتفسير النتائج أو استنباط جداول أخرى وفق مقاييس زمانية ومكانية أو بحسب فروض جديدة.

والتبويب والترتيب بالتعريف عبارة عن تنظيم البيانات في صورة أسطر أو أعمدة ضمن مجالات متجانسة مقرونة بمفرداتها الإحصائية التي تشترك بصفة معينة أو أكثر. وعادة التبويب يرتبط ارتباطاً وثيقاً بالمنهج المتبع وطرق التحليل والإمكانيات التقنية المتوافرة وطبيعة الظاهرة والبيانات والهدف المرجو من البحث.

9: تحليل البيانات الإحصائية :

يرتبط التحليل الإحصائي بالهدف وطبيعة البيانات وخصائصها والوصف الإحصائي للبيانات رقمياً أو بيانياً، يمهد تمهيداً سليماً للتحليل الإحصائي المعمق والمناسب لأنه يوضح الخصائص الإحصائية للظاهرة. وبقدر ما يشعر الباحث أنه كلما غالى في اختيار الطرق الإحصائية المتناهية الدقة أمكنه الوصول إلى نتائج دقيقة.

10: تفسير نتائج التحليل:

تعدّ عملية تفسير النتائج الإحصائية بالغة الأهمية في علم الإحصاء وأفضل تفسير تلك المعبر عنه بصفة رياضية ويجب ألا يتجاوز التفسير حدوده، وعلى الباحث أن يلتزم حدود نتائجه العلمية دون مبالغة أو زيادة حتى لا يضل القارئ في فهم نتائجه وبصورة عامة يجب أن يكون التفسير جواباً دقيقاً وصريحاً للمشكلة المقدّمة واختباراً دقيقاً لدقة الفروض المقدّمة في بداية التحليل الإحصائي.

11: القرار الإحصائي:

يبدأ التقرير الإحصائي من حيث بدأت المشكلة باختيارها وصياغتها وينتهي إلى حيث انتهت بالتحليل والتفسير النهائي، أي أنه يسجل خطوات البحث في تطورها خطوة تلو الخطوة ليكون بذلك أقرب إلى الموضوعية العلمية والتنظيم المنطقي ويشترط في القرار الإحصائي الوضوح والإيجاز والموضوعية مراحل أو خطوات البحث الإحصائي .

طرق عرض البيانات الاحصائية

إنَّ البيانات التي نحصل عليها قد تكون على أشكال مختلفة، فمنها على شكل قيم عددية مُفْردة، وبعضها الآخر قد يَعْرِض تَغيُّرَ ظاهرةٍ ما مع مرور الزمن أو مع مسمَّيات كـالبلدان، أو المدن، أو مع كليهما معاً، وبعضها الآخر قد يكون مجمَّعاً في جداول. لذلك سنبحث في تنظيم البيانات وفق اتجاهين:

الأول: يهتمّ بتنظيم البيانات المُفْرَدة التي تنتج مباشرة عن الدراسـة الإحصـائية (كميَّةً كانـت أم نوعيّة) في جداول تُدعى الجداول التكرارية، ومن ثمَّ تمثيل هذه البيانات في عروض بيانية مناسبة.

الثاني: يهتمّ بتجميع البيانات المُفْـرَدة الكميَّةً فقـط في جـداول مـن نـوع خـاص تُـدعى جـداول التوزيـع التكرارية، حيث يُقَال عن البيانات المقدَّمة بهذه الجداول إنَّهـا بيانـات مجمَّعـة (أو مبوَّبـة، أو مجدولـة)، ومن ثمَّ تمثيل بيانات هذه الجداول في عروض بيانية مناسبة.

وصولاً إلى الهدف المنشود ولذلك يجب ترتيبها أو تبويبها في مجالات بأسلوب معين يضمن معه المحافظة على البيانات دون حذف مخل بجوهرها

وفيما يلي أهم أنواع التبويب:

1- التبويب حسب مؤشر نوعي واحد: مثال تبويب السكان بحسب المهن أو الحالة التعليمية.

2- التبويب حسب مؤشر كمي واحد : مثال تبويب الطلاب بحسب معدلات النجاح أو بحسب الأطوال أو الأوزان.

3- التبويب حسب مؤشرين كميين: مثال توزع الطلاب في كلية بحسب الطول والوزن.

4- التبويب حسب مؤشرين نوعيين مثال توزع الطلاب حسب لون العيون والشعر.

5- التبويب حسب مؤشرين واحد كمي والآخر نوعي مثال توزع الطلاب حسب الطول والجنس.

6- التبويب التحليلي: وهو تبويب كمي يطبق على أحد مؤشرين مرتّبطين يمثل أحدهما المؤشر المتحول والآخر المؤشر التابع ويستخدم هذا النوع من التبويب ولإظهار العلاقات الإرتباطية بين المؤشرين

كليات الخليج
Gulf Colleges

7- التبويب الجغرافي وهو عبارة عن تبويب أو ترتيب البيانات الإحصائية حسب مكان حدوثها أو المتعلّقة بموقع مكاني جغرافي محدّد خلال فترة زمنية معينة.

8- التبويب الأبجدي عبارة عن ترتيب وتبويب البيانات الإحصائية المتعلّقة بظاهرة ما حسب الأحرف الأبجدية

البيانات الخام :

لدى تنفيذ دراسة إحصائية معيّنة حول ظاهرة ما وجمع البيانات حول هذه الظاهرة تنتج لدينا بيانات مُفْرَدة تُدعى بيانات خام.

إذا كان عدد البيانات صغيراً فإنَّه يمكن التعامل مع هذه البيانات بشكل مباشر (مع كل قيمة على انفراد) لدراستها، وأمّا إذا كان عدد البيانات كبيراً نسبياً، فإنَّه قد يكون من الصعب التعامل معها بشكل مباشر، ولذلك لا بدَّ من تقديم طرائق تسهل التعامل مع هذه البيانات كي يتمّ الاستفادة منها بأقصى

كليات الخليج
Gulf Colleges

طرق العرض الجدولى للبيانات الاحصائية

الجدول الإحصائي عبارة عن ترتيب منظم للبيانات الإحصائية في صورة صفوف وأعمدة بقصد إبراز أهمية تلك البيانات ولتسهيل المقارنة فيما بينها كهذا يعدّ إعداد الجدول الإحصائي من أهم مراحل تحليل البيانات الإحصائية لأنه مرتّبط ارتباطاً وثيقاً بطبيعة البيانات –الهدف المنشود- طبيعة الظاهرة –والتقنيات المستخدمة في التحليل الرياضي.

في حالة متغير كيفي نضع في هذه الخانات صفة المتغير.
في حالة متغير كمي منفصل نكتب قيم المتغير ويجب أن ترتب تصاعديا أو تنازليا.
في حالة متغير كمي متصل تكتب قيم المتغير على

المتغير X_i	التكرار المطلق n_i
x_1	n_1
x_2	n_2
.	.
.	.
.	.
x_k	n_k
المجموع $\sum$	$N = \sum n_i$

نضع في هذه الخانات عدد المفردات المقابلة لكل صفة أو قيمة أو فئة للمتغير الإحصائي

أنواع الجداول الإحصائية:

1- الجداول الإحصائية العامة:

وتشمل تلك الجداول التي تكتفي برصد البيانات فيها فقط دون الرغبة بتحليلها وتعدّ مرجعاً سهلاً للباحثين مثال ذلك المجموعات الإحصائية الرسمية أو المجموعات الإحصائية الخاصة (جداول التعداد جداول المواليد والوفيات والتقارير الإحصائية تقارير منظمة الصحة العالمية- نقابة الصيادلة شركات تصنيع الادوية

2-الجداول الإحصائية الخاصة:

وتشمل الجداول التي يشكلها الباحث، انطلاقاً من الجداول العامة السابقة الذكر بقصد إجراء بحث معين يبرز أهمية ظاهرة معينة في صورة مبسطة

3- الجداول البسيطة: وهي عبارة عن جداول تستخدم الترتيب أو التبويب لظاهرة معينة واحدة زماناً أو مكانا

جدول1 :توزع الطلاب في جامعة ما حسب الزمة الدموية خلال فترة زمنية معينة:

الزمرة الدموية	عدد الطلاب
O+	650
B	280
A	190
A-	175

- **الجداول المزدوجة**: عبارة عن جداول بسيطة تستخدم الترتيب أو تبويب البيانات لظاهرة معينة والتي يمكن تقسيمها إلى مجموعتين أو جزأين مع الزمان أو المكان.

كليات الخليج
Gulf Colleges

جدول 2 : توزع الطلاب الجامعة حسب الكليات والجنس مع الزمن:

الأعوام	كليات الصيدلة		كليات الطب البشري	
النوع	ذكور	إناث	ذكور	إناث
2021-2020	250	200	1020	580
2023-2022	400	300	1200	720
2024-2023	600	400	1300	850

كليات الخليج
Gulf Colleges

5- **الجداول المركبة** :عبارة عن جداول تستخدم لتبويب البيانات أو لتبويب مؤشر كمي واحد أو أكثر متفرع إلى مؤشرات جزئية أخرى

جدول3 : يبين عدد الطلاب في الجامعات السورية فرضي

الجامعات	عدد الطلاب / ألف طالب					
	مستجدون		طلاب قد		متخرجون	
	ذكور	إناث	ذكور	إناث	ذكور	إناث
دمشق	8115	8657	48853	43136	4584	4041
حلب	6573	4735	31138	178876	2765	1587
تشرين	3641	4737	14267	20581	1410	2014

6.-**الجداول التفضيلية**: عبارة عن جداول ترتب فيها المتغيّرات المدروسة وتعطى ترتيباً معيناً أو يقوم الباحث بإعطاء ترتيب معين.

جدول 4 : ذلك تفضيل مجموعة من الكلية لأقسام كلية الاقتصاد كما يلي:

الأقسام / الطلاب	صيدلة	طب بشري	طب اسان	ادارة مشافي	تمريض
AA	1	2	3	4	5
BA	5	2	4	3	1
CC	4	5	1	4	2

(مثال فرضي)

7-**جداول التشابه**: عبارة عن جداول ترتب فيها الموضوعات أو المتغيّرات حسب درجة قرابتها أو تشابهها مع بعضها البعض والتي تتراوح قيمتها بين الصفر والواحد. كما في المثال الآتي:

مثال جدول يبين التشابة بين فعالية بعض الادوية من شركتين (فرضي)

الشركة M_1 / الشركة M_2	A	B	C	D	E
A	0.3	0.7	0.8	0.1	0.9
B	0.1	0.9	0.2	0.4	0.6
C	0.2	0.4	0.8	0.6	0.7
D	0.6	0.8	0.3	0.9	0.2
E	0.4	0.5	0.7	0.5	0.6

كليات الخليج
Gulf Colleges

3- الشروط الواجب توفرها في الجدول الإحصائي:

1. رقم الجدول: يوضع لكل جدول رقم معين ويتم ترقيم الجداول إما بشكل.
2. عنوان الجدول: يجب أن تكون لكل جدول عنوان كامل ومختصر يفسر محتوياته
3. -عناوين الخانات: يجب أن تكون الخانات مرتّبة ومختصرة سواء للخانات الرأسية أو الأفقية
4. - خانة المجموع: يجب أن يتضمن الجدول خانة للمجموع الأفقي أو العمودي وللمجموع الإجمالي.
5. - الوحدات المستعملة يفضل تحديد الوحدات المستعملة في البيانات وتكتب عادة تحت عنوان الخانة
6. - مصدر البيانات إن تحديد مصدر البيانات أمر ضروري جداً يحدّد مدى الوثوق فيها
7. المذكرات التفسيرية توضع هذه المذكرات عادة تحت الجدول مباشرة وفوق المصدر على أن توضع برموز أو بحروف، وليس بأرقام منعاً للالتباس
8. ترتيب البيانات: يساعد ترتيب البيانات على سهولة فهم وتحليل بيانات الجدول وإمكانية المقارنة فيما بينها
9. البساطة يجب أن يتصف الجدول الإحصائي بالبساطة ويجب أن يكون الجدول الإحصائي مغلقاً
10. يفضل أن يكتب وترتب الجداول بنمط واحد في جميع أجزاء التقرير
11. تفضل الجداول الإحصائية المربعة الشكل قدر الإمكان.

شكراً لكم على حسن الاستماع

كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة الخامسة

طرق عرض البيانات الإحصائية النوعية:

1- عرض البيانات النوعية (الاسمية):

إن البيانات النوعية (données qualitives) والتي تشتمل على متغيّر واحد من المستوى الاسمي، يمكن تبويبها بسهولة ويسر، نظراً لأنها تكون مصنفة في أقسام متمايزة أثناء عملية جمعها

جدول (2) : تكرارات مجموعة من الأفراد بحسب متغيّر الحالة الاجتماعية

الحالة الاجتماعية	التكرارات
أعزب	30
متزوج	52
مطلق	7
أرمل	11
مجموع	100

إذا رغب الباحث المقارنة بين توزيعي مجموعتين مختلفتين في العدد الكلي لأفرادهما بالنسبة لمتغيّر اسمي معين، فعندئذ لا تجوز المقارنة باستخدام التكرارات وإنما ينبغي إيجاد النسب المئوية لتكرارات الأقسام لكل من المجموعتين ويحسب التكرار النسبي المئوي كما يلي:

كليات الخليج
Gulf Colleges

- التكرار النسبي = تكرار الفئة / مجموع التكرارات
- التكرار المئوي = التكرار النسبي x 100

مثال : جدول يبين تقديرات 60 طالب في الاحصاء 1

عدد مرات الإصابة	التعداد	التكرار	التكرار النسبي	التكرار المئوي
0	~~IIII~~ III	8	8/40 = 0.20	0.20×100 = 20 %
1	~~IIII~~ ~~IIII~~ ~~IIII~~ III	18	18/40 = 0.45	0.45×100 = 45 %
2	~~IIII~~ ~~IIII~~ II	12	12/40 = 0.30	0.30×100 = 30 %
3	II	2	2/40 = 0.05	0.05×100 = 5 %
Total		**40**	**1**	**100 %**

نفترض لدينا مجموعتين من الأفراد إحداهما تشتمل على 100 فرد والأخرى تشتمل على 200 فرد وإذا أدرنا المقارنة بين تكرارات أقسام متغيّر الحالة الاجتماعية لكل من المجموعتين لا بد من حساب التكرارات النسبية المئوية



ثانياً: الجدول التكراري النسبي والجدول التكراري المئوي للبيانات الوصفية:

نهتم كثيرًا بمعرفة التكرارات النسبية والتكرارات المئوية لكل فئة من فئات الجدول التكراري. فمعرفتنا بالتكرار النسبي للفئة أكثر فائدة من مجرد معرفة التكرار.

تعريف

نعرف التكرار النسبي والتكرار المئوي للفئة كما يلي:

أ– التكرار النسبي للفئة = تكرار الفئة ÷ عدد البيانات

= تكرار الفئة ÷ مجموع التكرارات

ب– التكرار المئوي للفترة = التكرار النسبي × 100%

مثال:

أوجد الجدول التكراري النسبي والجدول التكراري المئوي في مثال

الحل:

الفئة أو الصفة (التقدير)	التكرار (f)	التكرار النسبي= f/n	التكرار المئوي= (f/n) * 100%
A	6	6/60 = 0.100	10.0%
B	8	8/60 = 0.133	13.3%
C	16	16/60 = 0.267	26.7%
D	22	22/60 = 0.367	36.7%
E	8	8/60 = 0.133	13.3%
المجموع	$n = 60$	**1.000**	**100%**

التوزيعات التكرارية غير المبوبة

Ungrouped Data

التوزيع التكراري:

هو وسيلة لتنظيم وتجميع الدرجات أو القياسات أو البيانات بشكل منظم في مجموعات، ومن شأن هذا التنظيم أو التجميع تلخيص بيانات التوزيع في عدد محدود من هذه المجموعات لتسهيل معالجتها رياضياً. ولإنشاء جدول التوزيع التكراري غير المبوب ترتب الدرجات ترتيباً تنازلياً أو تصاعدياً ويسجل عدد مرات تكرار كل درجة منها.



- عرض البيانات عن طريق الجداول

التوزيع (الجدول) التكراري		
التكرار f	العلامات	الدرجة x
2	||	92
2	||	93
3	|||	94
6	𝍸 |	95
1	|	96
1	|	97
1	|	98
3	|||	99
1	|	100
$\sum f = 20$	مجموع التكرارات (الطلاب)	

التوزيع (الجدول) التكراري النسبي		
التكرار النسبي $\bar{f}$ $(f / \sum f =)$	التكرار f	الدرجة x
2/20 = 0.1 or 0.1 x 100 = 10%	2	92
2/20 = 0.1 or 0.1 x 100 = 10%	2	93
3/20 = 0.15 or 0.15 x 100 = 15%	3	94
6/20 = 0.3 or 0.3 x 100 = 30%	6	95
1/20 = 0.05 or 0.05 x 100 = 5%	1	96
1/20 = 0.05 or 0.05 x 100 = 5%	1	97
1/20 = 0.05 or 0.05 x 100 = 5%	1	98
3/20 = 0.15 or 0.15 x 100 = 15%	3	99
1/20 = 0.05 or 0.05 x 100 = 5%	1	100
$\sum \bar{f} = 1$ or $\sum \bar{f} = 100\%$	$\sum f = 20$	

جداول التوزيع التكراري لبيانات مبوّبة:
Grouped data

التبويب : عبارة عن مجموعات من المتغيّرات المتجانسة التي تتميز بحد أدنى وحد أعلى لمدى هذا التجانس أما التكرارات فهي عدد المفردات التي تندرج تحت فئة معينة من هذه الفئات

- خطوات إعداد جدول التوزيع التكراري:

1- ترتيب البيانات ترتيباً منتظماً تصاعدياً أو تنازلياً.

2- حساب المدى الكلي : وهو الفرق بين أكبر وأصغر قيمة في قيم التوزيع أي:

المدى الكلي = أكبر قيمة – أصغر قيمة

$R = X\,max - X\,mini$

3- حساب عدد المجالات (الفئات) ويتم عادة باستخدام إحدى المعادلتين التاليتين:

أ- معادلة سترجس وتعطى بالعلاقة التالية:

$$K=1+3.322 \log (n)$$

٤- تعيين مراكز الفئات Class Centers:

إنَّ مركز الفئة يساوي نصف مجموع حدّيها الأعلى والأدنى (أية فئة كانت العمليّة أو الفعليّة)، ويُنظر إلى هذه القيمة (أي إلى مركز الفئة) على أنها الممثّل لكل البيانات التي ستنتمي إلى هذه الفئة، ولذلك سوف نلاحظ أنَّ لهذه القيمة دوراً مهمّاً جداً لدى حساب القيم العددية المميّزة للبيانات المجمّعة في جداول توزيع تكرارية.

$$\text{مركز الفئة} = \frac{\text{الحد الأعلى} + \text{الحد الادنى}}{2}$$



مثال :

في أحد البحوث التي أجريت لدراسة مستوى الهيموجلوبين قام الباحث باختيار عينة مكونة من خمسين شخصا فحصل على البيانات التالية:

17.0	17.7	15.9	16.2	16.2	17.1	15.7	17.3
14.6	15.8	15.3	16.4	13.7	16.2	16.4	16.1
14.0	16.2	16.4	14.9	17.8	16.1	15.5	18.3
15.9	15.3	13.9	16.8	15.9	16.3	17.4	15.0
14.2	16.1	15.7	15.1	17.4	16.5	14.4	16.3
16.3	15.9	16.7	15.1	15.8	13.5	17.0	15.8
17.5	17.3						

أوجد جدول التوزيع التكراري لبيانات مستوى الهيموجلوبين لهؤلاء الأشخاص باستخدام الفترات التالية:

12.95–13.95, 13.95–14.95, 14.95–15.95,
15.95–16.95, 16.95–17.95, 17.95–18.95.

الحل:

أ– المتغير = مستوى الهيموجلوبين (متغير كمي).

ب– حجم العينة = عدد البيانات = n = 50.

ج– البيانات عبارة عن قيم مستوى الهيموجلوبين وهي بيانات كمية.

د– أكبر قيمة= 18.3 وأصغر قيمة = 13.5.

نقوم بتلخيص البيانات السابقة في توزيع تكراري وفق الخطوات التالية:

1. جدول تفريغ البيانات:

التكرار (f) (عدد الأشخاص)	العلامات	مستوى الهيموجلوبين (فترة الفئة)
3	|||	12.95 — 13.95
5	𝍸	13.9 5— 14.95
15	𝍸 𝍸 𝍸	14. 95— 15.95
16	𝍸 𝍸 𝍸		15. 95— 16.95
10	𝍸 𝍸	16. 95— 17.95
1			17. 95— 18.95

2. جدول التكراري لمستوى الهيموجلوبين:

التكرار (f) (عدد الأشخاص)	مستوى الهيموجلوبين (فترة الفئة)
3	12.95 — 13.95
5	13.9 5— 14.95
15	14. 95— 15.95
16	15. 95— 16.95
10	16. 95— 17.95
1	17. 95— 18.95
$n = 50$	المجموع

كليات الخليج
Gulf Colleges

مثال

لنفرض أن لدينا بيانات عددها n=45 ولنفرض أن هذه البيانات مسجلة إلى ثلاث خانات عشرية بحيث أن أصغر قيمة هي 2.517 وأكبر قيمة هي 2.548. والمطلوب هو تكوين الفترات التقريبية والفترات الفعلية بفرض أن الحد الأدنى للفترة التقريبية الأولى (الدنيا) هو $L_1 = 2.515$.

الحل:

من المعلومات المعطاة نستنتج ما يلي:

$$k = 1 + 3.322\log(45) = 6.49 \approx 7$$

$$R = X_{\max} - X_{\min} = 2.548 - 2.517 = 0.031$$

$$u = 0.001$$

$$w = \frac{R}{k} = \frac{0.031}{7} = 0.004443 \approx 0.005$$

باستخدام القيمة المعطاة للحد الأدنى للفترة التقريبية الأولى (الدنيا) $L_1 = 2.515$ ، وباستخدام القوانين التالية:

$$U_i = L_i + 0.005 - 0.001 \ ; i = 1, 2, \ldots, 7$$

$$L_i = U_{i-1} + 0.001 \ ; \ i = 2, 3, \ldots, 7$$

$$L_i^* = L_i - \frac{0.001}{2} = L_i - 0.0005 ; \ \ i = 1, 2, \ldots, 7$$

$$U_i^* = U_i + \frac{0.001}{2} = U_i + 0.0005 ; \ \ i = 1, 2, \ldots, 7$$



نوجد الفترات التقريبية والفترات الفعلية كما في الجدول التالي:

الفترات التقريبية		الفترات الفعلية	
الحد الأعلى	الحد الأدنى	الحد الأعلى	الحد الأدنى
2.519	2.515	2.5195	2.5145
2.524	2.520	2.5245	2.5195
2.529	2.525	2.5295	2.5245
2.534	2.530	2.5345	2.5295
2.539	2.535	2.5395	2.5345
2.544	2.540	2.5445	2.5395
2.549	2.545	2.5495	2.5445

نلاحظ أن أصغر قيمة 2.517 تقع في الفترة الأولى (الدنيا)، وأكبر قيمة 2.548 تقع في الفترة الأخيرة (العليا). كما نلاحظ أن الفترات التقريبية غير متلاصقة (الفرق بين كل فترة والتي تليها مقدار وحدة دقة 0.001)، وأما الفترات الفعلية فهي متلاصقة.

MANARA UNIVERSITY

(2−1−4) الجدول التكراري النسبي والجدول التكراري المئوي للبيانات الكمية:

كما مر معنا سابقًا، فإننا نعرف التكرار النسبي والتكرار المئوي للفئة كما يلي:

أ− التكرار النسبي للفئة = تكرار الفئة ÷ عدد البيانات

= تكرار الفئة ÷ مجموع التكرارات

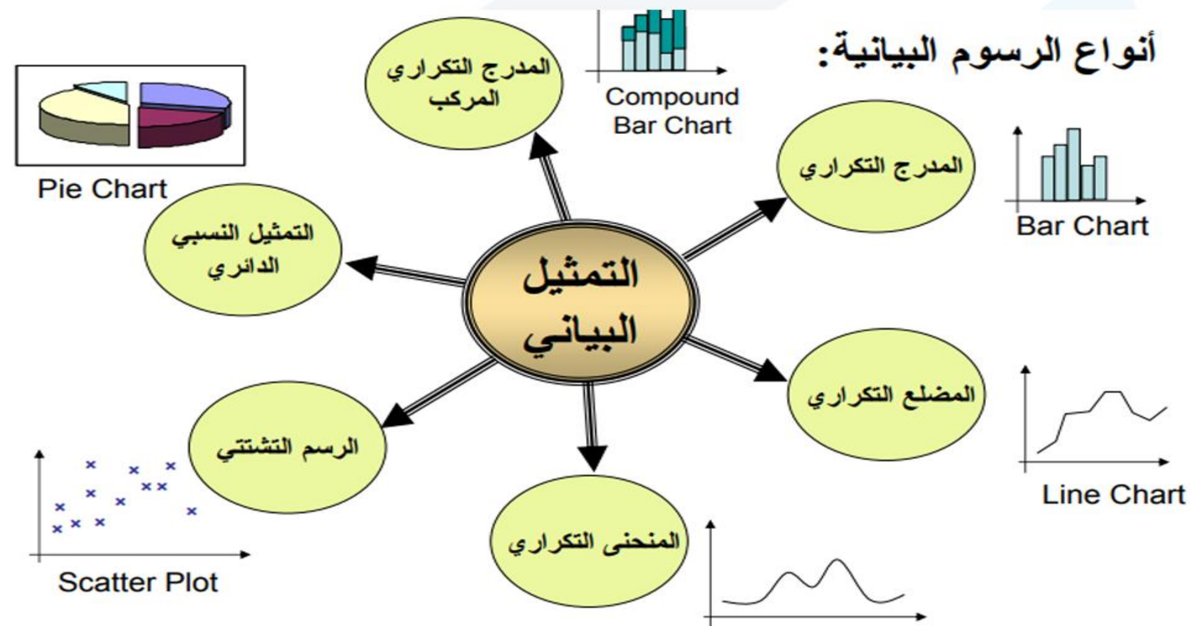
ب− التكرار المئوي للفترة = التكرار النسبي × 100%

مثال:

أوجد الجدول التكراري النسبي والجدول التكراري المئوي في مثال (.)

الحل:

مستوى الهيموجلوبين (فترة الفئة)	التكرار (f)	التكرار النسبي= f/n	التكرار المئوي= (f/n) * 100%
12.95 — 13.95	3	3/50= 0.06	6%
13.9 5— 14.95	5	5/50= 0.10	10%
14. 95— 15.95	15	15/50= 0.30	30%
15. 95— 16.95	16	16/50= 0.32	32%
16. 95— 17.95	10	10/50= 0.20	20%
17. 95— 18.95	1	1/50= 0.02	2%
المجموع	$n = 50$	1.00	100%


طرق العرض البياني للبيانات الإحصائية
Graphic Presentation
أنواع الرسوم البيانية:
التمثيل البياني
المدرج التكراري المركب
Compound Bar Chart
المدرج التكراري
Bar Chart
المضلع التكراري
Line Chart
المنحنى التكراري
الرسم التشتتي
Scatter Plot
التمثيل النسبي الدائري
Pie Chart

مقدّمة:

العرض البياني (Graphic Presentation) عبارة عن تمثيل للقياسات أو البيانات المتعلّقة بظاهرة معينة / نوعية – كمية/ بأشكال هندسية مرسومة حسب مقاييس معينة، لتوضيح سلوك هذه الظاهرة أو الصفة بالنسبة إلى عاملين مختلفين، **بهدف إبراز بعض خصائص هذه الظاهرة أو لمقارنة قيم ظاهرة ما حسب المكان أو تطورها بحسب الزمن.**

وأهم الشروط الواجب مراعاتها حتى تكون الأشكال الهندسية ذات فائدة أهمها:

a. **أن يكون لكل شكل بياني مهما كانت طبيعته عنو اناً واضحاً ومختصراً.**
b. **أن يعطى لكل شكل بياني رقم محدّد وفق طريقة ترقيم الأشكال المعتمدة.**
c. **تحميل الظواهر أو رموزها على محاور الإحداثيات.**
d. **يجب أن يحدّد وحدة القياس المستخدمة وكذلك توضيح المقياس المستخدم.**

2-3 : طرق العرض البياني للبيانات الإحصائية:

a- الأعمدة البيانية البسيطة : Bar Charts :

عبارة عن مستطيلات ترسم بعرض واحد وبأطوال مختلفة، حيث يدل الارتفاع على كمية البيانات المراد دراستها، وعادة يترك فواصل فيما بينها بحيث لا يتجاوز عرض الفاصل بين كل مستطيلين عرض قاعدة المستطيل الواحد تستخدم هذه الأعمدة في الحالات التالية:

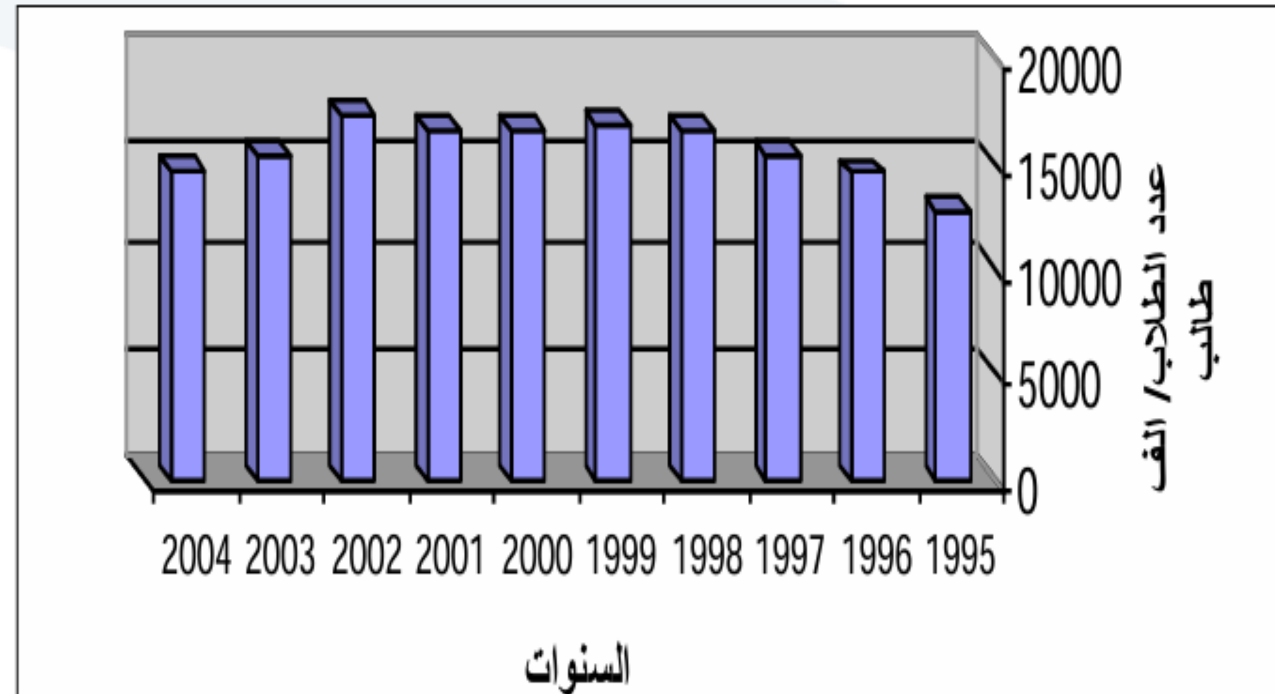
كليات الخليج
Gulf Colleges

- **المقارنة بين قيم ظاهرة واحدة حسب المكان أو تطوره حسب الزمن.**
- **المقارنة بين صفين أو أكثر.**
- **مقارنة مكونات الظاهرة في مكانين مختلفين أو فترتين مختلفتين.**

جدول يبين أعداد الطلاب خريجي الجامعات السورية من 1995 ولغاية 2004

العام	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
عدد الطلاب	12800	14600	15400	16700	17000	16650	16755	17462	15500	14800

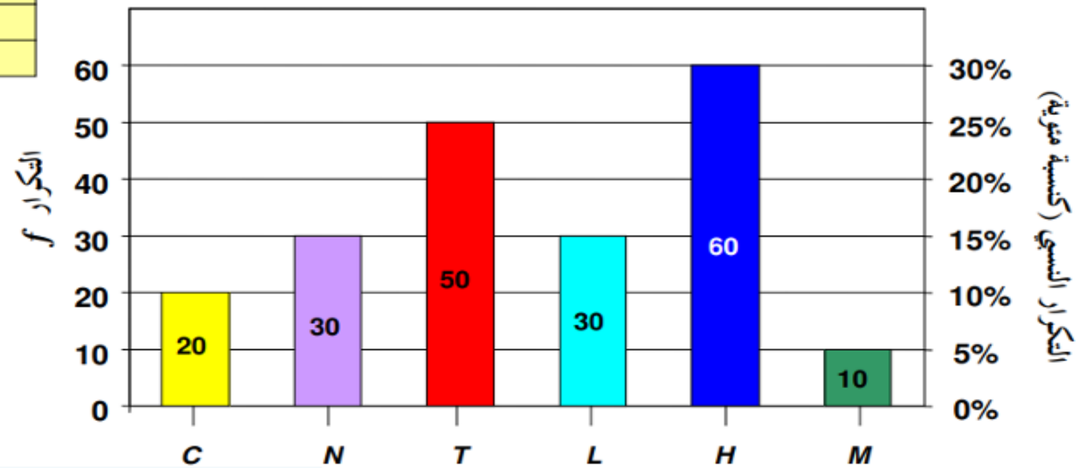
المطلوب : تمثيل هذه البيانات بواسطة الأعمدة البسيطة.


20000
15000
10000
5000
0
عدد الطلاب/ الف طالب
2004 2003 2002 2001 2000 1999 1998 1997 1996 1995
السنوات

شكل2: تمثيل البيانات بواسطة الاعمدة البيانية البسيطة

طريقة الأعمدة البسيطة

x	التكرار	التكرار النسبي
C	20	10%
N	30	15%
T	50	25%
L	30	15%
H	60	30%
M	10	5%



كليات الخليج
Gulf Colleges

b- الأعمدة البيانية المزدوجة:

تستخدم إذا كان الهدف من الرسم هو مقارنة بين ظاهرتين أو أكثر ولعدة سنوات، أو إذا كان لدينا بيانات مزدوجة بخواص مختلفة،

يبين عدد الطلاب في قسم الاقتصاد، بحسب الجنس من عام 1998 ولغاية 2004

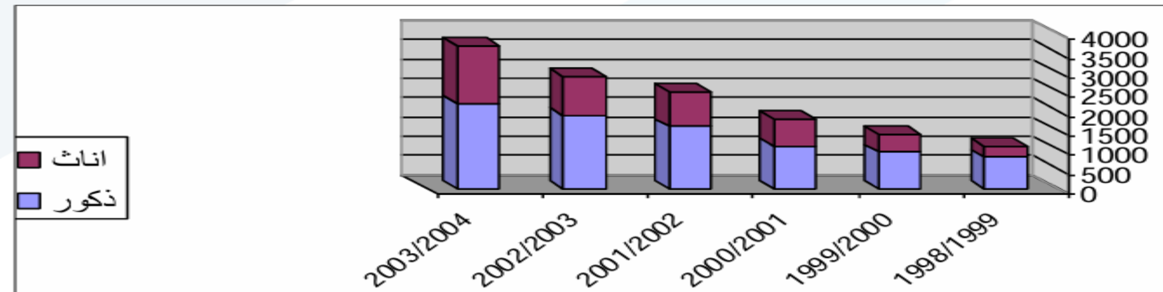
السنة	1998	1999	2000	2001	2002	2003	2004	مجموع
ذكور	800	950	1100	900	800	920	1200	
إناث	300	460	700	1200	1350	1450	1600	

ـ الأعمدة البيانية المقسمة (المجزأة):

تستخدم في نفس الحالات التي تستخدم فيها الأعمدة البيانية المزدوجة، ويتم الحصول عليها برسم عمود واحد يمثل جملة الظواهر محل الدراسة في كل سنة، كما في حالة الأعمدة البسيطة ثم نقسم كل عمود إلى مكوناته حيث يقاس كل جزء مع العدد الذي يمثله ونميز بين الأجزاء بالتظليل أو الألوان.

مثال لنعود إلى معطيات الأعمدة البيانية المزدوجة، والمطلوب تمثيلها سابقاً.

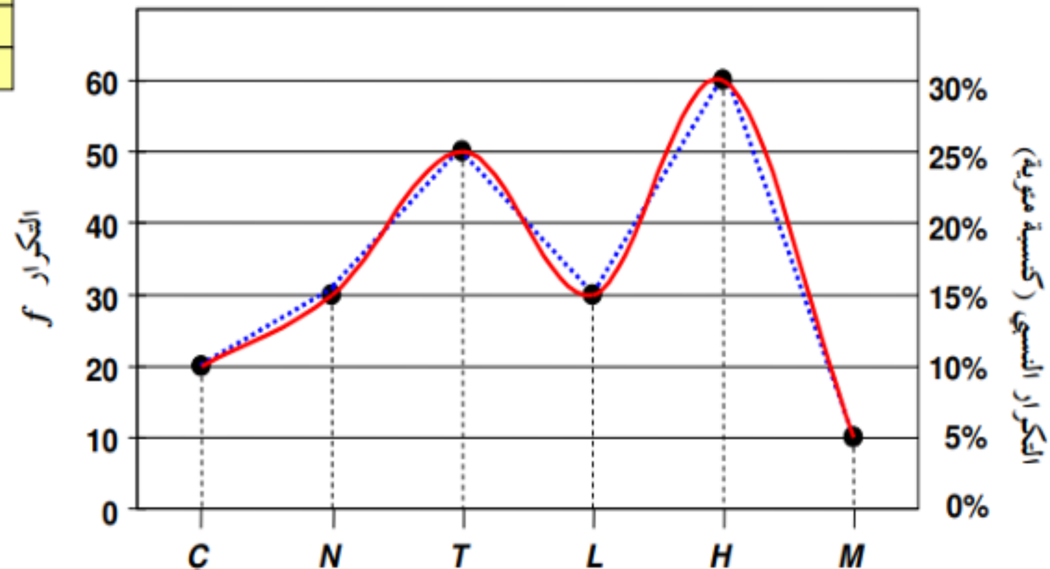
السنوات	1999/1998	2000/1999	2001/2000	2002/2001	2003/2002	2004/2003
ذكور	800	950	1100	1600	1900	2200
إناث	300	460	700	900	1000	1500



شكل 4: تمثيل عدد الطلاب بواسطة الأعمدة البيانية المقسمة. بحسب النوع.

طريقة المضلع أو المنحنى التكراري (التكراري النسبي)

x	التكرار	التكرار النسبي
C	20	10%
N	30	15%
T	50	25%
L	30	15%
H	60	30%
M	10	5%



-3-2 المجسمات البيانية (الأشكال المساحية):

a- الدوائر البيانية:

تستخدم الدوائر أو أنصافها في كثير من الحالات في التمثيل البياني للبيانات الإحصائية، ويتم عرض البيانات بهذا الأسلوب إما بواسطة دوائر مستقلة أو تمثل البيانات على سطح دائرة واحدة، حيث يعزز الأسلوب الأخير تقسيم الدائرة أو نصف الدائرة وحتى ربع الدائرة إلى عدة أجزاء

-عندما يراد تمثيل البيانات على سطح دائرة واحدة فقط، يتم وفق الآتي: إن مساحة الدائرة تساوي 360 وهي تقابل النسبة المئوية 100% أي أن مساحة القطاع المخصص للفئة أ تقابل زاوية مقدارها I/1 درجة ونسبتها المئوية p% وبذلك يمكننا حساب الزاوية المقابلة للنسب المختلفة من العلاقة التالية:

كليات الخليج
Gulf Colleges

$$yi = \frac{360^\circ \times pi\,\%}{100} = \frac{360^\circ \times 32.98}{100} = 118.7^\circ$$

أو الزاوية = $\frac{\text{عدد طلاب كل ك ل ية}}{\text{ال طلاب الإجمال ي ل ل ك ل يات}}$ × 360 ْ

أوعلى سطح نصف دائرة:

$$yi = \frac{180^\circ \times pi\,\%}{100} = \frac{180^\circ \times 32.98}{100} = 59.36^\circ$$

أو الزاوية = 180 ْX $\frac{\text{عـــدد طـــلاب كـــل كليـــة}}{\text{عـــــدد الطـــــلاب الإجمـــــالي للكليـــــات}}$

أوعلى سطح ربع دائرة

$$yi = \frac{90^\circ \times pi\,\%}{100} = \frac{90^\circ \times 32.98}{100} = 29.68^\circ$$

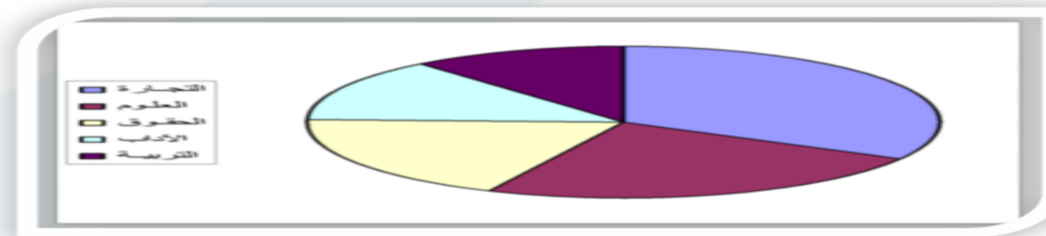
الزاوية = $\frac{\text{عدد طلاب كل ك ل ية}}{\text{ال طلاب الإجمال ي ل ل ك ل يات}}$

مثال عدد طلاب بعض الكليات لعام2000 والمطلوب تمثيلها بواسطة دائرة بيانية واحدة.

الكليات	العدد	Pi%	$Y^{\circ}=360$	$Y^{\circ}=180$	$Y^{\circ}=90$
التجارة	12500	23.98	118.7	59.56	29.66
العلوم	9100	24	86.4	43.2	21.6
الحقوق	7000	18.70	64	33.66	16.83
الآداب	5100	13.6	48.45	24.22	12.11
التربية	4200	11	39.6	19.8	9.9
المجموع	37900	%100	360	180	90

المصدر : فرضي

ويمكن رسمها بيانياً وفق الأشكال التالية



- التمثيل البياني للتوزيعات التكرارية:

- طرق عرض بيانات التوزيعات التكرارية:

1- المدرج التكراري.
2- المضلع التكراري.
3- المنحني التكراري.
4- المدرج التكراري التجميعي الصاعد والهابط.
5- المنحني التكراري التجميعي الصاعد والهابط.

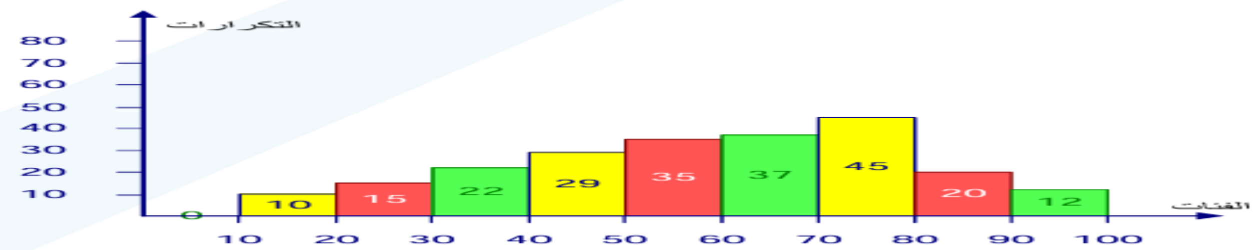
1- المدرج التكراري <u>The histogram</u>

يستخدم المدرج التكراري في تمثيل توزيع الدرجات أو القيم وذلك برسم شكل بياني على هيئة مستطيلات متلاصقة لأن المتغيّر متصل ومن المستوى الفئوي أو النسبي وعدد المستطيلات يساوي عدد فئات التوزيع، وقاعدة كل منها هي الجزء الذي يمثل طول الفئة، وارتفاعه يمثل التكرار، والمساحة الكلية للمستطيلات تتناسب مع التكرار الكلي للتوزيع ولرسم المدرج التكراري، نرسم محورين متعامدين يمثل المحور الأفقي (السيني) فئات الدرجات، والخط أو المحور العمودي (الصادي) يمثل التكرارات (n_i).

يبين الجدول التوزيع التكراري لدرجات225 طالباً في الإحصاء الحيوي

الفئات	التكرارات n_i
20-10	10
30-20	15
40-30	22
50-40	29
60-50	35
70-60	37
80-70	45
90-80	20
100-90	12
مجموع التكرارات	225

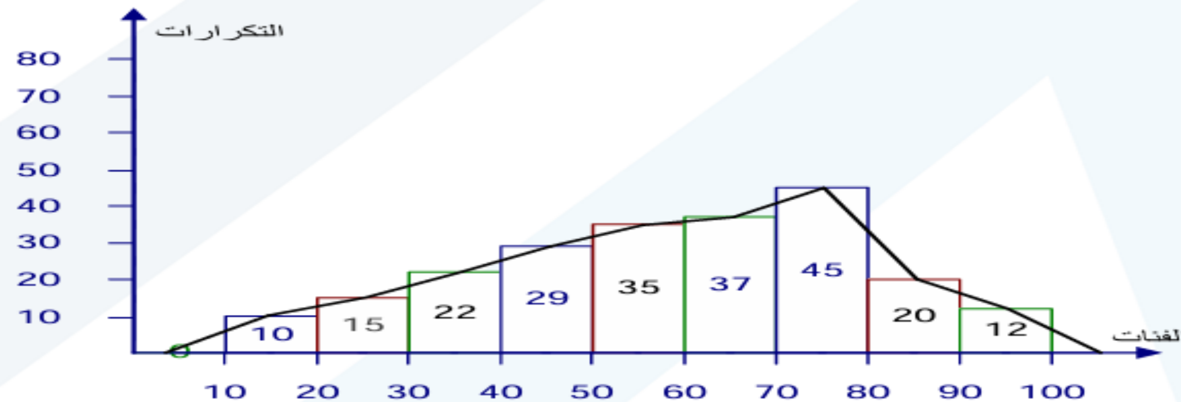
المطلوب: رسم المدرج التكراري لهذه الدرجات.



شكل 8: المدرج التكراري لدرجات 225 طالباً في الإحصاء

- المضلع التكراري: Polygon de frequene :

نجد أن تكرار كل فئة مركز في منتصف الفئة وهذا هو الفرق بين المدرج التكراري والمضلع التكراري. ولرسم المضلع التكراري نقوم برسم محورين متعامدين كما سبق في حالة المدرج التكراري ولكن يجب هنا أن نضيف فئتين إحداهما تسبق الفئة الدنيا والأخرى تعقب الفئة العليا. وتعدّ تكرار كل منها صفر، ثم نعين نقاط تناظر تكرار كل فئة (بما في ذلك الفئتان اللتان تكرار كل منها صفر) فوق منتصف كل فئة، وبعدها نصل بين هذه النقاط بخط منكسر، المضلع التكراري هو الخط المنكسر الواصل بين منتصفات القواعد العليا للمدرج التكراري والممتد من إحدى ناحيته إلى منتصف الفئة السابقة للتوزيع .ومن الناحية الأخرى إلى منتصف الفئة التي تعقب فئات التوزيع. وبذلك يكون المضلع التكراري مقفلاً وتكون مساحته مساوية بالضبط لمساحة المدرج التكراري. مثال:



كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة السادسة

مقاييس النزعة المركزية

تستخدم مقاييس النزعة المركزية ومقاييس التشتت لوصف خصائص البيانات

فمقاييس النزعة المركزية تصف مدى تمركز البيانات حول قيمة معينة مما يسمح باستخدام هذه القيمة المركزية لتمثل البيانات ومقاييس النزعة المركزية أنواع منها الوسط الحسابي والوسيط والمنوال وأي منهم يفضل على الآخر في حالات، وبصورة عامة المقياس يكون جيد اذا توفرت فيه كل أو معظم الصفات التالية

أ- أن يأخذ كل المشاهدات في الاعتبار
ب- اذا كان سهل الحساب
ت- ان يكون قابل للحساب الجبري
ث- ان لا يتأثر بالقيم المتطرفة

الوسط الحسابي

- يعرف أيضا بالمتوسط، ويرمز له ب $\overline{X}$ تقرأ x بار

الوسط الحسابي من المفردات:-

- اذا كان لدينا المفردات التالية

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad ....x_n$

فإن الوسط الحسابي لهذه المفردات = مجموع المفردات / عدد المفردات

$$\overline{X} = \frac{\sum x}{n}$$

حيث

$\sum x$ تمثل مجموع المشاهدات أو المفردات

n تمثل عدد المشاهدات

يستخدم الوسط الحسابي في المعايرة كما يستخدم في مقارنة المجموعات اذا كانت متجانسة

كليات الخليج
Gulf Colleges

مثال

أحسب الوسط الحسابي للبيانات التالية:

1- 21 14 11 8 6

2- 6 19 7 24 22 11 9

الحل

$$\overline{X} = \frac{\sum x}{n}$$

-1

$$\frac{6+8+11+14+21}{5} = \frac{60}{5} = 12$$

-2

$$\frac{9+11+22+24+7+19+6}{7} = \frac{98}{7} = 14$$

كليات الخليج
Gulf Colleges

مثال

أحسب الوسط الحسابي للبيانات التالية:

1- 4 5 7 3 8 9

2- 5 7 3 8 9 100

ماذا تلاحظ

الحل

1-
$$\overline{X} = \frac{\sum x}{n}$$

$$\frac{9+8+3+7+5+4}{6} = \frac{36}{6} = 6$$

2-
$$\frac{9+8+3+7+5+100}{6} = \frac{132}{6} = 22$$

نلاحظ التغير الكبير في الوسط الحسابي بسبب تغير مفردة واحدة وتعتبر هذه المفردة بعيدة عن بقية المفردات أي شاذة عنها أي ان الوسط الحسابي تأثر بالقيمة الشاذة.

تنبيهات

- لا يشترط ان تكون قيمة الوسط الحسابي موجودة ضمن المشاهدات او المفردات فقد تكون موجودة أو غير موجودة ولكن يجب أن تكون قيمة الوسط الحسابي محصورة بين أصغر وأكبر مفردة
- الفرق بين أي مفردة والوسط الحسابي يسمى بانحراف المفردة عن الوسط ومجموع انحرافات المفردات عن الوسط دائما يساوي صفر أي أن

$$\sum (x_i - \bar{x}) = 0$$

الوسط الحسابي المرجح

في بعض الاحيان تكون هنالك مفردات او مشاهدات أكثر أهمية من غيرها فمثلاً اذا كان لدينا المفردات التالية

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad \ldots . x_m$

وكان لكل منها وزن معين كما يلي

$w_1 \quad w_2 \quad w_3 \quad w_4 \quad w_5 \quad \ldots . w_m$ فان الوسط الحسابي يحسب بالقانون

$$\overline{X} = \frac{\sum_{i=1}^{m} w_i x_i}{\sum_{i=1}^{m} w_i}$$

الوسط الحسابي من الجداول التكرارية

يحسب الوسط الحسابي من الجداول التكرارية أو ما يعرف بالبيانات المبوبة بالقاعدة التالية

$$\overline{X} = \frac{\sum fx}{\sum f}$$

حيث

$\sum fx$ تمثل مجموع حاصل ضرب التكرارات في مراكز الفئات

$\sum f$ تمثل مجموع التكرارات

x تمثل مراكز الفئات

يعتمد هذا القانون على الجداول المنتظمة فقط اما في حالة الغير منتظمة يفضل استخدام التكرار المعدل

مثال

البيانات التالية توضح التوزيع العمري لأسرة مكونة من خمسة عشر فرداً أحسب متوسط عمر العائلة

الفئات العمرية c	-1	-5	-9	-13	21-17
عدد الافراد f	3	4	3	4	1

الحل

الوسط الحسابي من الجداول يحسب بالقانون

$$\overline{X}=\frac{\sum fx}{\sum f}$$

f.x	x	f	c
9	3	3	-1
28	7	4	-5
33	11	3	-9
60	15	4	-13
19	19	1	21- 17
149		15	المجموع

$$\overline{X}=\frac{149}{15}=9.933$$ سنة

مثال

البيانات التالية تمثل انتاج مصنع أسمنت الجنوب خلال 18 شهر حيث الانتاج بآلاف الاطنان أحسب متوسط الانتاج

كمية الانتاج	الشهور
-1	2
-5	4
-9	6
-13	4
21-17	2

الحل

الوسط الحسابي من الجداول يحسب بالقانون

f.x	x	f	c
6	3	2	-1
28	7	4	-5
66	11	6	-9
60	15	4	-13
38	19	2	21-17
198		18	المجموع

طن

$$\overline{X} = \frac{\sum fx}{\sum f}$$

$$\overline{X} = \frac{198}{18} = 11$$

مثال

أحسب متوسط البيانات التالية

f	c
5	6-2
4	11-7
6	16-12
3	21-17
2	22-فأكثر

الحل

يجب أولاً اغلاق الجدول بأخذ طول الفئة السابقة حتى نستطيع حساب الوسط الحسابي

$$\overline{X} = \frac{\sum fx}{\sum f}$$

f.x	x	f	c
20	4	5	6-2
36	9	4	11-7
84	14	6	16-12
57	19	3	21-17
48	24	2	26-22
245		20	المجموع

$$\overline{X} = \frac{245}{20} = 12.25$$

مزايا وعيوب الوسط الحسابي

أولاً مزايا الوسط:

- يأخذ كل المشاهدات في الاعتبار
- سهل الحساب
- يفضل الوسط في حالة البيانات المعتدلة وفي حالة عدم وجود قيم متطرفة

ثانياً عيوب الوسط:

- يتأثر بالقيم أو المشاهدات الشاذة أو المتطرفة ويتأثر بعدد الدرجات ويميل للاستقرار كلما كان هذا العدد كبير (100فأكثر)
- لا يمكن حسابه من البيانات النوعية
- لا يمكن حسابه في حالة الجداول المفتوحة

ملحوظة: هنالك طرق اخرى لحساب الوسط الحسابي منها طريقة الوسط الفرضي ومنها طريقة الوسط الفرضي والعامل المشترك

واجب

- أحسب متوسط البيانات التالية

1- 6 8 2 4 6 12 15 12

2-

f	c
5	8-2
8	15-9
6	22-16
3	29-23
2	36-30

واجب

1- البيانات التالية تمثل الانتاج بمئات الأطنان لمزارع بمنطقة الباحة

5 7 8 32 12 4 9 15 6 13

حيث تكررت هذه الانتاجية في 100 مزرعة كما يلي وعلى الترتيب

12 10 11 6 11 14 9 8 14 5

أحسب متوسط الإنتاج

2- البيانات التالية تمثل انتاج أحد المصانع بمدينة جدة خلال 24 شهر حيث الانتاج بآلاف الوحدات أحسب متوسط الانتاج

f	c
5	8-2
8	15-9
6	22-16
3	29-23
2	36-30

كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة السابعة

الوسيط

- هو القيمة التي تتوسط القيم

الوسيط من المفردات : اذا كان لدينا المفردات التالية

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad ....x_n$

لحساب الوسيط يجب أولاً ترتيب المفردات تصاعديا او تنازليا، فإذا كان عدد المفردات(n) فردي فإن رتبة الوسيط = $\frac{1 + n}{2}$

الوسيط هو المفردة التي تقع وسط المفردات بعد الترتيب

أما اذا كان عدد المفردات زوجي فإنه توجد مفردتين في الوسط رتبة المفردة الأولى 2/n ورتبة الثانية $\frac{n}{2} + 1$ أي التي تليها

والوسيط في هذه الحالة = متوسط المفردتين (مجموع المفردتين على 2) عادة نرمز للوسيط ب M

مثال

- أحسب الوسيط فيما يلي:

1- 6 9 3 7 2

2- 100 6 9 3 2

3- 6 24 18 8 12 4

4- 12 15 12 6 4 2 8 6

5- مقبول جيد راسب ممتاز جيد جدا

الحل

1- أولاً الترتيب 2 3 6 7 9 عددها فردي ثانياً رتبة الوسيط $\frac{n+1}{2}$ $\left(\frac{5+1}{2}\right)$ المفردة الثالثة ثالثا الوسيط 6

2- الترتيب 2 3 | 6 9 100 رتبة الوسيط $(n+1)/2$ $(5+1)/2$ المفردة الثالثة اذن الوسيط 6 لاحظ ان الوسيط لم يتأثر بالقيمة الشاذة 100

3- الترتيب 24 18 12 8 6 4 عددها زوجي رتبة المفردة الاولى $(n/2)$ أي الثالثة ورتبة المفردة الثانية $(n/2 +1)$ أي التي تليها الرابعة الوسيط = $\left(\frac{8+12}{2}\right) = 10$

كليات الخليج
Gulf Colleges

تابع للحل

4- الترتيب 15 12 12 8 6 6 4 2 عددها
زوجي رتبة المفردة الاولى (n/2) أي الرابعة و التي تليها
الخامسة اذن الوسيط = (6+8/2) = 7

5- الترتيب راسب مقبول جيد جيد جدا ممتاز

عددها فردي رتبة الوسيط $\frac{n+1}{2}$ $\left(\frac{5+1}{2}\right)$ المفردة الثالثة

اذن الوسيط جيد

الوسيط من البيانات المبوبة

يحسب الوسيط من الجداول التكرارية بالقاعدة

$$M = L + \frac{(\frac{n}{2} - c_1)}{(c_2 - c_1)} * h$$

L تمثل الحد الأدنى للفئة الوسيطة

h تمثل طول الفئة الوسيطة

c_1 التكرار المتجمع الصاعد السابق للفئة الوسيطة

c_2 التكرار المتجمع الصاعد اللاحق للفئة الوسيطة

$\sum f = n$

مثال

البيانات التالية تمثل الدرجات التي أحرزها 100 طالب في امتحان للفيزياء أحسب وسيط الدرجات

-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	c
1	4	13	17	21	18	15	7	3	1	f

الحل

نكون التكراري الصاعد

ترتيب الوسيط هو

n/2 = 100/2= 50

L = 60   h = 5

C	التكرار الصاعد f.d
أقل من 40	0
أقل من 45	1
أقل من 50	5
أقل من 55	18
أقل من 60	35
أقل من 65	56
أقل من 70	74
أقل من 75	89
أقل من 80	96
أقل من 85	99
أقل من 90	100

n/2=50

الفئة الوسيطة 60-65

تابع للحل

- من الجدول الصاعد نجد ان $c_2 = 56$ $c_1 = 35$

الوسيط هو:

$$M = L + \frac{(\frac{n}{2} - c_1)}{(c_2 - c_1)} \times h$$

$$M = 60 + \frac{(50 - 35)}{(56 - 35)} \times 5 = 63.57$$ درجة

كليات الخليج
Gulf Colleges

Gulf Colleges كليات الخليج

مثال

البيانات التالية تمثل التوزيع العمري لمجتمع مكون من100 فرد أحسب وسيط العمر لهذا المجتمع

الفئات العمرية c	20-	25-	30-	35-	40-	45-	50-	55-
عدد الافراد f	6	14	16	17	14	19	8	6

الحل

نكون التكراري الصاعد

ترتيب الوسيط هو

$n/2 = 100/2 = 50$

$L = 35 \quad h = 5$

C	التكرار الصاعد f.d
أقل من 20	0
أقل من 25	6
أقل من 30	20
أقل من 35	36
أقل من 40	53
أقل من 45	67
أقل من 50	86
أقل من 55	94
أقل من 60	100

n/2=50

الفئة الوسيطة 35-40

تابع للحل

- من الجدول الصاعد نجد ان $c_2 = 53$ $c_1 = 36$

الوسيط هو:

$$M = L + \frac{(\frac{n}{2} - c_1)}{(c_2 - c_1)} \times h$$

$$M = 35 + \frac{(50-36)}{(53-36)} \times 5 = 39.11$$ سنة

كليات الخليج
Gulf Colleges

مثال

أحسب الوسيط لبيانات الجدول التالي والتي تمثل الدخل الشهري ل 20 موظف بالجامعة حيث الدخل بآلاف الريالات

f	c
5	-2
4	-7
6	-12
3	-17
2	22-فأكثر

الحل

نكون التكراري الصاعد

ترتيب الوسيط هو

$n/2 = 20/2 = 10$

$L = 12 \quad h = 5$

الفئة الوسيطة 12-17

التكرار الصاعد f.d	C
0	أقل من 2
5	أقل من 7
9	أقل من 12
15	أقل من 17
18	أقل من 22
20	أقل من 27

n/2=10

تابع للحل

- من الجدول الصاعد نجد ان $c_2 = 15$ $c_1 = 9$
الوسيط هو:

$$M = L + \frac{(\frac{n}{2} - c_1)}{(c_2 - c_1)} \times h$$

$$M = 12 + \frac{(10-9)}{(15-9)} \times 5 = 12.833$$

وسيط الدخل = 12833 ريال

كليات الخليج
Gulf Colleges

الوسيط من العرض البياني

يمكن ايجاد الوسيط من العرض البياني من خلال نقطة تقاطع المنحنى الصاعد مع المنحنى الهابط

مثال:مستخدما بيانات الجدول التالي وضح الوسيط بيانيا

f	c
12	-20
16	-25
30	-30
25	-35
12	-40
5	-45

الحل

نكون الجدول الصاعد والهابط/

الحدود العليا	التكرار الصاعد fd
أقل من 20	0
أقل من 25	12
أقل من 30	28
أقل من 35	58
أقل من 40	83
أقل من 45	95
أقل من 50	100

الوسيلة ← n/2 →

الحدود العليا	التكرار الهابط
20 فأكثر	100
25 فأكثر	88
30 فأكثر	72
35 فأكثر	42
40 فأكثر	17
45 فأكثر	5
50 فأكثر	0

الفئة الوسيطة
n/2 →

مزايا وعيوب الوسيط

أولا مزايا الوسيط:

- لا يتأثر بالقيم الشاذة
- يمكن حسابه من البيانات النوعية اذا أمكن ترتيبها وكان عددها فردي
- يمكن حسابه من الجداول التكرارية المفتوحة
- يفضل استخدام الوسيط للمقارنة والمعايرة وخصوصا عندما يكون التوزيع ملتويا

ثانيا العيوب:

- لا يأخذ كل القيم في الاعتبار
- لا يمكن حسابه من البيانات النوعية اذا كان عددها زوجي

كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة الثامنة

كليات الخليج
Gulf Colleges

المنوال

المنوال من المفردات: المنوال هو القيمة الأكثر شيوعا أو الأكثر تكراراً ويستخدم المنوال كثير جدا في مجال التسويق

مثال: حدد المنوال فيما يلي

1- 6 2 8 6 1 2 2

2- 3 8 9 8 5 3 7

3- 1 9 11 3 8 7 12

4- أحمد حمد حسين محمد حسن محمد

الحل

1- المنوال هو 2
2- يوجد منوالان هما 3 8
3- لا يوجد منوال
4- المنوال هو محمد

المنوال من الجداول التكرارية

يحسب المنوال والذي يرمز له ب D من البيانات المبوبة بالقاعدة التالية

$$D = L + \frac{\Delta_1}{(\Delta_1 + \Delta_2)} \times h$$

حيث :

L الحد الادنى للفئة المنوالية

h طول الفئة المنوالية

Δ_1 الفرق بين أكبر تكرار والتكرار السابق له

Δ_2 الفرق بين أكبر تكرار والتكرار اللاحق له

تسمى هذه الطريقة بطريقة بيرسون

مثال

احسب المنوال لبيانات الجدول التالي والتي تمثل الكميات المعروضة والكميات المباعة لمنتجات احدى الشركات في 5مدن حيث الكميات بآلاف الوحدات

المعروضة	المباعة
2-	1
7-	4
12-	8
17-	5
22-فأكثر	4

كليات الخليج
Gulf Colleges

الحل

- الفئة المنوالية هي الفئة المناظرة لأكبر تكرار (الفئة الثالثة)

$\Delta_2 = 8-5=3$ $\Delta_1 = 8 -4=4$

h=5 L = 12

$$D = L + \frac{\Delta_1}{(\Delta_1 + \Delta_2)} \times h$$

$$D = 12 + \frac{4}{(4+3)} \times 5 = 14.857$$

كليات الخليج
Gulf Colleges

مثال

احسب المنوال لبيانات الجدول التالي

f	c
12	-4
4	-10
8	-16
5	-22
5	28-فأكثر

كليات الخليج
Gulf Colleges

الحل

- الفئة المنوالية هي الفئة المناظرة لأكبر تكرار (الفئة الأولى)

$\Delta_2 = 12 - 4 = 8 \quad \Delta_1 = 12 - 0 = 12$

h=6 L = 4

$$D = L + \frac{\Delta_1}{(\Delta_1 + \Delta_2)} \times h$$

$$D = 4 + \frac{12}{(12 + 8)} \times 6 = 7.6$$

مثال

احسب المنوال لبيانات الجدول التالي

f	C
6	-2
8	-8
15	-12
5	-17
4	24 -22

الحل

الجدول غير منتظم نحسب التكرار المعدل

f d f/r	r طول الفئة	f	C
1	6	6	-2
2	4	8	-8
3	5	15	-12
1	5	5	-17
2	2	4	24 -22

الحل

- الفئة المنوالية هي الفئة المناظرة لأكبر تكرار في التكرار المعدل (الفئة الثالثة)

$\Delta_2 = 3 - 1 = 2$ $\Delta_1 = 3 - 2 = 1$

$h = 17 - 12 = 5$ $L = 12$

$$D = L + \frac{\Delta_1}{(\Delta_1 + \Delta_2)} \times h$$

$$D = 12 + \frac{1}{(1+2)} \times 5 = 13.667$$

مزايا وعيوب المنوال

أولاً المزايا

- سهل الحساب
- لا يتأثر بالقيم المتطرفة أو الشاذة
- يمكن حسابه من البيانات النوعية
- يمكن حسابه من الجداول التكرارية المفتوحة

ثانياً العيوب

- لا يأخذ كل المشاهدات في الاعتبار
- قد يكون هنالك منوال أو أكثر من منوال وقد لا يوجد منوال

العلاقة بين الوسط والوسيط والمنوال

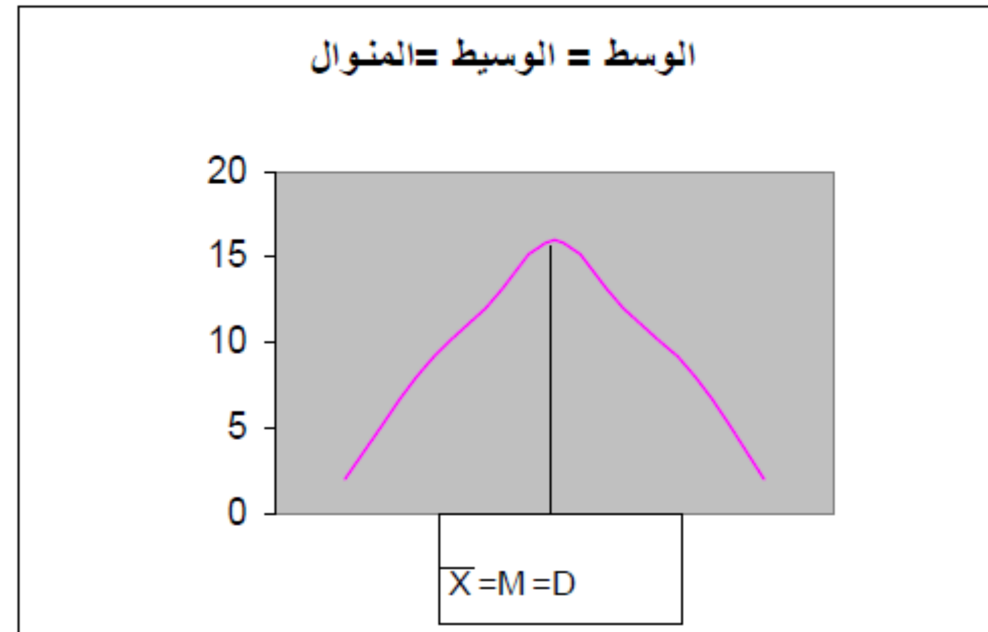
اذا كان التوزيع متماثلا فان $\bar{X} = M = D$

وإذا كان ملتويا جهة اليمين فان $\bar{X} > M > D$ اما وإذا كان ملتويا جهة اليسار فان $\bar{X} < M < D$

في حالة التوزيعات التكرارية القريبة من التماثل فإن العلاقة تعطى ب الوسط الحسابي- المنوال=3(الوسط الحسابي – الوسيط)

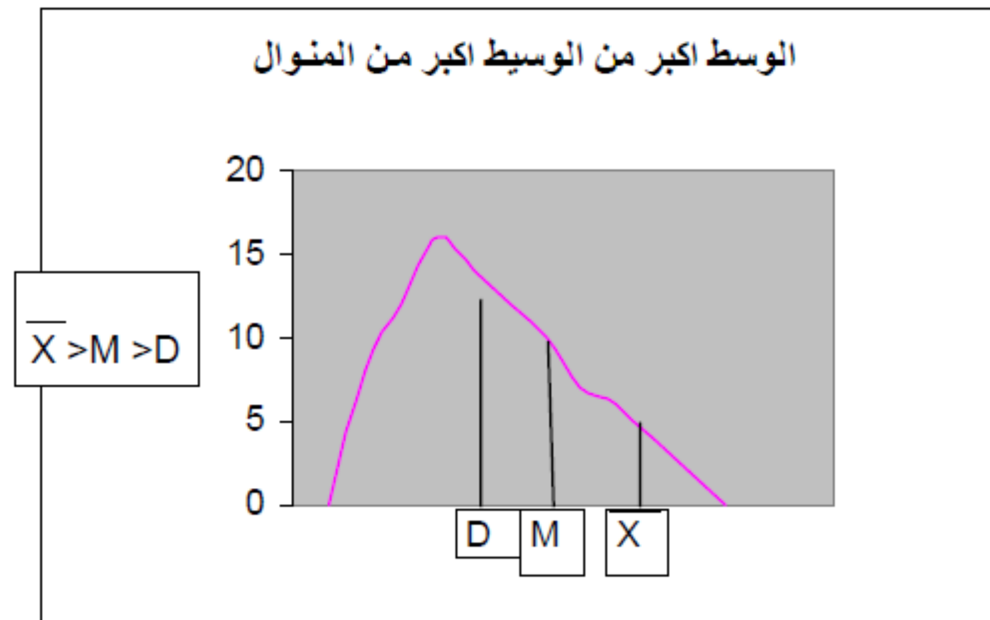
العلاقة بين المتوسطات من خلال الرسم

- في حالة التماثل الوسط = الوسيط = المنوال



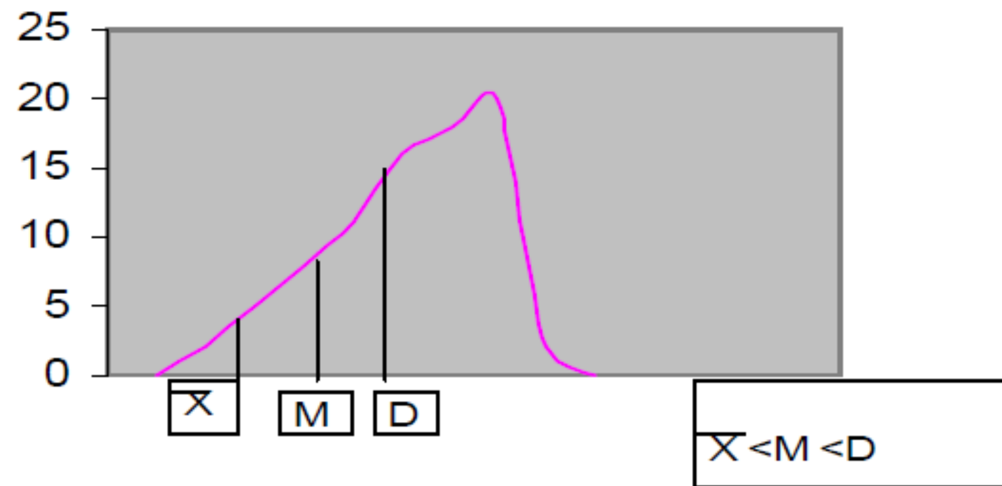
العلاقة بين المتوسطات من خلال الرسم

- التواء نحو اليمين



التواء نحو اليسار

في هذه الحاله يكون الوسط اصغر المقاييس الثلاثة



مثال

اذا كان لدينا توزيع قريب من التوزيع المتماثل وكان وسطه الحسابي 27 ومنواله 31 فما قيمة وسيط هذا التوزيع

الحل

بما أن التوزيع قريب من التوزيع المتماثل فإن

الوسط الحسابي – المنوال = 3(الوسط الحسابي- الوسيط)

27 – 31 = 3(27 – الوسيط)

-4 = 81 – 3الوسيط   -85 = -3 الوسيط

الوسيط =28.333

كليات الخليج
Gulf Colleges

مثال

اذا كان X متغير يتبع لتوزيع قريب جدا من التوزيع المتماثل
أحسب الوسط الحسابي ل x اذا كان وسيطة 12 ومنواله 6

كليات الخليج
Gulf Colleges

الحل

الوسط الحسابي – المنوال = 3(الوسط الحسابي- الوسيط)

- المنوال =2 الوسط الحسابي – 3 الوسيط

- 6 = 2 الوسط الحسابي – 36

الوسط الحسابي ل x = 15

كليات الخليج
Gulf Colleges

الحل

الوسط الحسابي – المنوال = 3(الوسط الحسابي- الوسيط)

- المنوال =2 الوسط الحسابي – 3 الوسيط

- 6 = 2 الوسط الحسابي – 36

الوسط الحسابي ل x = 15

كليات الخليج
Gulf Colleges

واجب

احسب الوسط والوسيط والمنوال للتوزيعات التالية ثم علق على شكل المنحنى

f	c
5	- 10
20	-20
15	- 30
10	-40
5	-50

f	c
5	-10
10	-20
25	- 30
10	-40
5	-50

f	c
5	-10
10	-20
15	- 30
20	-40
5	-50

كليات الخليج
Gulf Colleges

الاحصاء الحيوي

Biostatistics



المحاضرة التاسعة

مقاييس التشتت

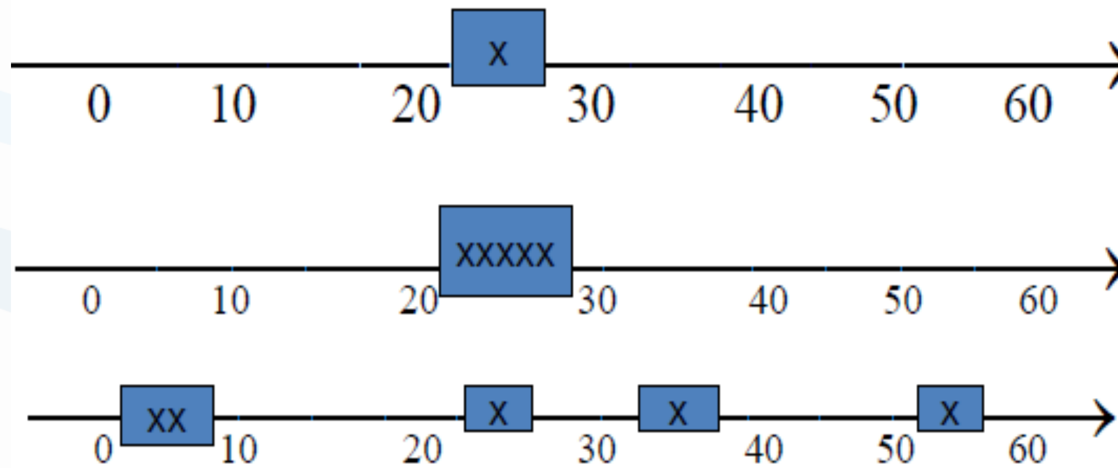
- في كثير من الأحيان نجد مقاييس النزعة المركزية غير كافية لوصف البيانات فمثلاً اذا كان لدينا المجموعات التالية من المفردات

الاولى	24	24	24	24	24
الثانية	20	21	24	26	29
الثالثة	3	8	24	33	52

فإذا حسبنا متوسط المجموعات = 24 وإذا حسبنا وسيط المجموعات ايضا 24 ولكن واضح أن هنالك فرق من حيث مدى تقارب مفردات هذه المجموعات وتباعدها عن بعضها

توضيح

يمكن توضيح ذلك كما يلي



الأولى

الثانية

الثالثة

المدى

يعطي فكرة سريعة عن مدى تغير الظاهرة ويرمز له ب R

- المدى من المفردات:

المدى = أكبر مفردة – أصغر مفردة

مثال اذا كان درجات الطلاب في واجب للإحصاء كما يلي

6 8 11 15 16 12 18 5 14 20

أحسب المدى الحل

$R = 20 - 5 = 15$

كليات الخليج
Gulf Colleges

واجب

احسب المدى فيما يلي

1-	24	24	24	24	24
2-	20	21	24	26	29
3-	3	8	24	33	52

المدى من البيانات المبوبة

يحسب بطريقتين

الاولى عن طريق الفئات:

المدى = الحد الاعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

الثانية عن طريق مراكز الفئات:

المدى = مركز الفئة الأخيرة – مركز الفئة الاولى

الحل

المدى = الحد الأعلى للفئة الأخيرة – الحد الادنى للفئة الأولى

22 -12= 10

f	c
3	-12
2	-14
3	-16
4	-18
3	22 -20

كليات الخليج
Gulf Colleges

مزايا وعيوب المدى

أولا المزايا

- سهولة حسابه
- مقياس يعطي فكرة سريعة عن تفاوت البيانات

ثانياً العيوب

- لا يأخذ كل القيم في الاعتبار وقد تكون احداهما شاذة
- يصعب حسابه في البيانات الوصفية
- يصعب حسابه من الجداول المفتوحة

كليات الخليج
Gulf Colleges

الانحراف المتوسط

يحسب الانحراف المتوسط من المفردات بالقاعدة

$$M\,D = \frac{\sum |x - \bar{x}|}{n}$$

حيث n تمثل عدد المفردات وx تمثل المفردات

كليات الخليج
Gulf Colleges

مثال

- أحسب الانحراف المتوسط للبيانات التالية

-1	24	24	24	24	24
-2	20	21	24	26	29
-3	3	8	24	33	52

الحل

- نحسب الوسط الحسابي

$$\overline{X} = \frac{\sum x}{n}$$

1- $$\overline{X} = \frac{120}{5} = 24 \qquad MD = \frac{\sum |x - \overline{x}|}{n} = \frac{0}{5} = 0$$

2- $$\overline{X} = \frac{120}{5} = 24 \qquad MD = \frac{|20-24| + |21-24| + |24-24| + |26-24| + |29-24|}{5} =$$

$$MD = \frac{4+3+0+2+5}{5} = \frac{14}{5} = 2.8$$

3- $$\overline{X} = \frac{120}{5} = 24 \qquad MD = \frac{|3-24| + |8-24| + |24-24| + |33-24| + |52-24|}{5} =$$

$$MD = \frac{21+16+0+9+28}{5} = \frac{74}{5} = 14.8$$

التباين والانحراف المعياري

هو متوسط مربعات الانحرافات عن الوسط ويختلف التباين للمجتمع والذي يرمز له ب σ^2 وتقرأ سيجما تربيع عن تباين العينة s^2

التباين من المفردات في حالة المجتمع

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

في حالة العينات

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

اما الانحراف المعياري فهو الجذر التربيعي للتباين

مثال

اختيرت عينة من 10 مكالمات دولية فكان طول المكالمات بالدقائق كما يلي

10 6 12 15 20 8 4 9 3 13

أحسب التباين والانحراف المعياري لطول المكالمات

الحل

$$\bar{X} = \frac{\sum x}{n} = \frac{13+3+9+4+8+20+15+12+6+10}{10} = \frac{100}{10} = 10$$



تابع للحل

نحسب تباين العينة

$$S^2 = \frac{\sum (X_i - \overline{X})^2}{n-1}$$

$$s^2 = \frac{(10-10)^2 + (6-10)^2 + (12-10)^2 + (15-10)^2 + (20-10)^2 + (8-10)^2 + (4-10)^2 + (9-10)^2 + (3-10)^2 + (13-10)^2}{9}$$

$$s^2 = \frac{244}{9} = 27.111$$

الانحراف المعياري =

$$s = \sqrt{27.111} = 5.207 \text{ دقائق}$$

مثال

مجتمع مكون من 10 مفردات أحسب التباين والانحراف المعياري للمجتمع

10 8 12 13 8 15 7 12 16 9

الحل

$$\mu = \frac{\sum x}{N} = \frac{9+16+12+7+8+15+13+12+8+10}{10} = \frac{110}{10} = 11$$

كليات الخليج
Gulf Colleges

شكرا على حسن الاستماع